Collective modes in asymmetric nuclear matter

M. Colonna a, M. Di Toro a, A.B. Larionov a,b

a Laboratorio Nazionale del Sud, Via S. Sofia 44, I-95123 Catania, Italy
and University of Catania, Catania, Italy
b RRC “I.V. Kurchatov Institute”, Moscow 123182, Russia

Received 10 October 1997; revised 25 February 1998
Editor: J.-P. Blaizot

Abstract

The collective response of asymmetric nuclear matter is studied on the basis of the corresponding generalized Landau Dispersion Relations. Isoscalar-isovector coupling, disappearance of collectivity and possibility of new instabilities are discussed with accent on their relation to properties of the symmetry term in the nuclear equation of state. Implications for collective modes in $\beta$-unstable nuclei, for reaction mechanisms with radioactive heavy-ion beams and for nuclear astrophysics are presented. © 1998 Published by Elsevier Science B.V. All rights reserved.

PACS: 21.65.+f; 21.60.Ev; 24.30.Cz

Keywords: Asymmetry; Nuclear matter; Collective modes; Linear response; Landau damping; Instabilities

Nuclei far from $\beta$-stability represent a quite exciting research field not only for nuclear structure, testing new features of the nucleon-nucleon interaction, but also for dynamical effects of a large charge asymmetry. In order to have a general picture of new collective modes and of their relations to nuclear interaction properties, in this letter we follow a dispersion relation approach to study the small-amplitude collective response of asymmetric (neutron rich) nuclear matter in a large range of densities and for quite different effective interactions. In this way one can study collective vibrations as well as the onset of instabilities.

Important effects, not much depending on the particular used interaction, of charge asymmetry on the properties of collective motions will be presented. At variance, we will show that other results, like the existence of neutron-proton separation instabilities [1,2], crucially depend on the structure of the symmetry term in the nuclear equation of state. This is an open problem in the understanding of supernova explosions and neutron star formation [3,4].

Our results around normal density can be used as general guidelines in predicting the behaviour of volume collective modes in finite $\beta$-unstable nuclei. Collective motions around the neutron drip line have been studied in a RPA approach [5–7] revealing some interesting new effects: i) A strong coupling between isovector and isoscalar collective motions; ii) A quite large spreading of the strength in the low frequency region, only partially due to threshold contributions from loosely bound neutrons. Some first data with radioactive beams seem also to confirm the coupling between isoscalar and isovector strength [8].

At sub-normal densities we can foresee new sce-
narios for a dynamical formation of fragments with particular isotopic contents in collisions with radioactive beams.

We will follow a mean field approach, i.e. we study the propagation of pure Landau zero sound modes. This choice allows us to isolate charge asymmetry effects. The Landau dispersion relation (d.r.) [9–12] can be extended to the study of collective modes in asymmetric nuclear matter. To our knowledge, this extension was done for the first time in Ref. [13] where results were analysed at normal density only the zero-order Landau parameters.\(^1\)

However we have performed numerical calculations taking into account also first-order Landau parameters.

In the variation of the neutron (proton) mean field potential both isoscalar and isovector density variations will appear:

\[
\delta U_q = \alpha_q(\rho_n, \rho_p) \delta \rho_n + \beta_q(\rho_n, \rho_p) \delta \rho_p,
\]

where \(\rho_n \equiv \rho_n + \rho_p\) and \(\rho_q \equiv \rho_n - \rho_p\) are respectively isoscalar and isovector densities, \(n\) and \(p\) stand for neutrons and protons, and \(q = n, p, \tau_q = +1 (q = n), -1 (q = p)\). As a consequence, the linearized Vlasov equations for the isoscalar and isovector variations of the distribution function, \(\delta f_q = \delta f_n + \delta f_p\) and \(\delta f_i = \delta f_n - \delta f_p\), will be coupled:

\[
\begin{align*}
\frac{\partial \delta f_i}{\partial t} + \frac{p}{m} \frac{\partial \delta f_i}{\partial r} &- \left( \frac{\alpha_i}{\rho_n} + \frac{\beta_i}{\rho_p} \right) \frac{\partial \delta f_i}{\partial p} \delta \rho_n + \left( \frac{\alpha_i}{\rho_n} - \frac{\beta_i}{\rho_p} \right) \frac{\partial \delta f_i}{\partial p} \delta \rho_p = 0, \\
\frac{\partial \delta f_i}{\partial t} + \frac{p}{m} \frac{\partial \delta f_i}{\partial r} &- \left( \frac{\alpha_i}{\rho_n} - \frac{\beta_i}{\rho_p} \right) \frac{\partial \delta f_i}{\partial p} \delta \rho_n + \left( \frac{\alpha_i}{\rho_n} + \frac{\beta_i}{\rho_p} \right) \frac{\partial \delta f_i}{\partial p} \delta \rho_p = 0,
\end{align*}
\]

where

\[
f_i(\epsilon^q_p) = \frac{2}{(2\pi\hbar)^3} \Theta(\epsilon^q_n - \epsilon^q_p), \quad q = n, p
\]

is the unperturbed distribution function; \(\epsilon^q_n, \epsilon^q_p = (p^q)^2/(2m)\) and \(\epsilon^q_n = p^q/(2m)\) are respectively Fermi momentum, Fermi energy and kinetic energy of the nucleons of the sort. \(m = 938\) MeV is the bare nucleon mass.

In the general case of nuclear matter at arbitrary asymmetry \(I = (N - Z)/A\), starting from Eqs. (2), (3) we can derive, using the Landau procedure, a system of coupled equations for the \(k\)-component of the isovector and isoscalar density variation:

\[
\begin{align*}
\rho_{k, i, s} &= \sum_k \rho_k \exp(ik \cdot r - i\omega_k t) \\
\rho_k^i \left(1 + F_n^i L_n + F_p^i L_p\right) + \rho_k^s \left(F_n^i L_n - F_p^i L_p\right) &= 0, \\
\rho_k^i \left(F_n^i L_n - F_p^i L_p\right) + \rho_k^s \left(1 + F_n^i L_n + F_p^i L_p\right) &= 0.
\end{align*}
\]

where \(F_n^i\) and \(F_p^i\) are dimensionless constants:

\[
\begin{align*}
F_n^i &= \frac{3}{2} \rho_n \frac{p^{i(p)}_n}{\epsilon_i(p)} \alpha_i(p), \\
F_p^i &= \frac{3}{2} \rho_p \frac{p^{i(p)}_p}{\epsilon_i(p)} \beta_i(p),
\end{align*}
\]

\(\rho_i(p)\) is the Fermi momentum and \(L_n, L_p\) are Lindhard functions:

\[
L_{i(p)} \equiv L(s_{i(p)}); \quad L(s) = 1 - \frac{s}{2} \ln \left(\frac{s + 1}{s - 1}\right)
\]

with \(s_{i(p)} = \omega_i/k_F(u)^{(p)}\) being the zero sound phase velocity in neutron (proton) Fermi velocity units. The constants (6) can be expressed in terms of the usual zero-order Landau parameters as it follows:

\[
\begin{align*}
F_n^i &= \frac{1}{2} (F_{0 n}^{n(n)} + F_{0 p}^{n(p)}), \\
F_p^i &= \frac{1}{2} (F_{0 n}^{p(p)} + F_{0 p}^{p(p)}).
\end{align*}
\]

The d.r. will result by imposing the determinant of the system (5) to be zero:

\[
\begin{align*}
(1 + F_n^i L_n + F_p^i L_p)(1 + F_n^i L_n + F_p^i L_p) &= (F_n^i L_n - F_p^i L_p)(F_n^i L_n - F_p^i L_p), \\
F_n^i &= \frac{1}{2} (F_{0 n}^{n(p)} - F_{0 p}^{n(p)}).
\end{align*}
\]

Using the standard Landau parameters, relation (10) can be rewritten under the form:

\[
\begin{align*}
1 + F_{0 n}^{n(n)} L_n + F_{0 p}^{n(p)} L_p &+ (F_{0 n}^{n(p)} - F_{0 p}^{n(p)}) L_n L_p = 0.
\end{align*}
\]

\(^1\) However we have performed numerical calculations taking into account also first-order Landau parameters.
In the case of symmetric nuclear matter, we get back to the uncoupled d.r. for isoscalar and isovector modes [2,12]:
\[ 1 + F_0 L(s) = 0, \quad 1 + F_0' L(s) = 0, \]
\[ s = \omega/(k v_F^s) \]  \hspace{1cm} (12)
where \( F_0 = F_0^{nn} + F_0^{np} \) and \( F_0' = F_0^{nn} - F_0^{np} \) are the usual isoscalar and isovector Landau parameters [12].

Before discussing the solutions of the d.r. (10), it is instructive to make explicit the asymmetry dependence of variables \( s_n \) and \( s_p \). Since \( \rho_n = \rho_s(1 + I)/2, \rho_p = \rho_s(1 - I)/2 \), we have:
\[ s_n = s \left( \frac{p_F}{p_F^0} \right)^{1/3} = \frac{s}{(1 + I)^{1/3}} \]  \hspace{1cm} (13)
\[ s_p = s \left( \frac{p_F}{p_F^0} \right)^{1/3} = \frac{s}{(1 - I)^{1/3}}. \]  \hspace{1cm} (14)

Eqs. (13), (14) are particularly significant. Indeed, corresponding to solutions with \( 0 < s - 1 < 1 \) obtained for symmetric matter, in the asymmetric case we will see a 'diving' into the \( s_n, p < 1 \) region, i.e. the disappearance of the corresponding collective mode due to Landau damping. Physically it means that with increasing asymmetry the collective mode starts to interact with nucleons of an increasing Fermi sea (with neutrons, if \( I > 0 \), or protons, if \( I < 0 \)), until completely swallowed [13,15].

The above discussion is confirmed by solving numerically the d.r. (10). We have performed calculations with a Skyrme-like effective interaction:
\[ v_{12} = t_0 (1 + x_0 P_\rho) \delta(r_1 - r_2) \]
\[ + \frac{1}{6} t_3 (1 + x_3 P_\rho) \left[ \rho_s \left( \frac{r_1 + r_2}{2} \right) \right]^\gamma \delta(r_1 - r_2), \]  \hspace{1cm} (15)
with \( t_0 = -2973 \) MeV \cdot fm\(^4\), \( t_3 = 19034 \) MeV \cdot fm\(^{\gamma + 1}\), \( x_0 = 0.025 \) \( x_3 = 0 \) \( \gamma = 1/6 \). This parameterization gives the correct values for saturation density and binding energy of symmetric nuclear matter (see Fig. 1a). The compressibility modulus results equal to \( K = 201 \) MeV. This parameterization can be considered as a simplified SKM interaction [16] (we put \( m^* = m \) instead of \( m^* = 0.79 m \), as in the SKM interaction) and gives a good reproduction of the energy per nucleon in neutron matter, as calculated by Friedman and Pandharipande [17], in a wide range of densities. In the following we will refer to this interaction as the Soft interaction.

The potential energy density can be expressed under the form:
\[ \mathcal{H}_{\rho_{01}}(\rho_s, \rho_i) = \frac{A}{2} \frac{\rho_0^2}{\rho_0^4} + \frac{B}{\gamma + 2} \frac{\rho_0^{\gamma + 2}}{\rho_0^4} \]
\[ + \frac{C(\rho_i)}{2} \rho_i^2, \]  \hspace{1cm} (16)
where \( \rho_0 = 0.16 \) fm\(^3\) is the saturation density of nuclear matter and the coefficients \( A, B \) and \( C \) in the energy density functional (16) are connected to the Skyrme parameters as follows:
\[ A = \frac{3}{4} t_0 \rho_0, \quad B = \frac{\gamma + 2}{16} t_3 \rho_0^{\gamma + 1}, \]  \hspace{1cm} (17)
\[ C(\rho_i) = -\rho_0 \left[ \frac{t_0}{2} \left( x_0 + \frac{1}{2} \right) + \frac{t_3}{12} \left( x_3 + \frac{1}{2} \right) \rho_i^2 \right], \]  \hspace{1cm} (18)
Neutron and proton mean field potentials are:

\[
U_q = \frac{\delta \mathcal{H}_{pot}}{\delta \rho_q} = A \left( \frac{\rho_q}{\rho_0} \right) + B \left( \frac{\rho_q}{\rho_0} \right)^{y+1} + C \left( \frac{\rho_q}{\rho_0} \right) \tau_q + \frac{1}{2} \frac{dC}{d\rho_q} \frac{\rho_q^2}{\rho_0}, \tag{19}
\]

The coefficient \(C\) in Eqs. (16), (19) at saturation density is related to the symmetry coefficient in the Weissacker mass formula:

\[
a_{sym} = \frac{1}{2} \frac{\alpha^2}{\beta^2} \left( \frac{E}{A} \right)_{I=0} = \frac{\epsilon_F}{3} + C \left( \rho_0 \right) / 2 \tag{20}
\]

that with the parameterization considered and \(\epsilon_F = 37\) MeV, gives the standard value of the mass formula: \(a_{sym} = 28\) MeV.

The coefficients \(\alpha\) and \(\beta\) in Eq. (1) can be expressed as:

\[
\alpha_q = \frac{A}{\rho_0} + \left( \sigma + 1 \right) B \frac{\rho_q^y}{\rho_0^{y+1}} + \tau_q \frac{dC}{d\rho_q} \frac{\rho_q}{\rho_0} + \frac{1}{2} \frac{d^2C}{d\rho_q^2} \frac{\rho_q^2}{\rho_0}, \tag{21}
\]

\[
\beta_q = \frac{C}{\rho_0} + \tau_q \frac{dC}{d\rho_q} \frac{\rho_q}{\rho_0}. \tag{22}
\]

In the case of nuclear matter at small asymmetry \(I\), the coupling between Eqs. (2) and (3) vanishes, because \(\epsilon_I^p = \epsilon_I^s\), so \(\alpha_s = \alpha_p\) and \(\beta_s = \beta_p\) (see Eqs. (21), (22)). However from the above relations we can expect some coupling in heavy \(\beta\)-stable elements, particularly due to surface effects, where we can have larger values of \(dC/d\rho_q\) and \(\rho_q\).

The study of collective modes in asymmetric nuclear matter can be afforded also in the framework of momentum-dependent interactions, as in Ref. [13]. When introducing the effective mass \(m^*\), the formalism becomes more complicated since the Landau parameters \(F_{L=1}\) have to be considered and the solution of the dispersion relation is obtained by imposing the determinant of a matrix \((4 \times 4)\) to be zero. Therefore here we will skip the formalism, but we will present also some results obtained using momentum dependent interactions in order to show that the most relevant effects of the charge asymmetry on the properties of collective modes are quite robust and do not depend very much on the interaction considered. We have done calculations using the interaction BPAL32 (see Eq. (9) and Tables 1, 2 in Ref. [18]), that can be considered as the extension to asymmetric cases of the momentum-dependent GBD interaction [19].

Fig. 1a shows the density behaviour of the energy per nucleon \(E/A\) obtained at \(I = 0\) and \(I = 1\). It is possible to observe a similarity of the two interactions in the case of symmetric nuclear matter. Indeed the BPAL32 interaction has a compressibility modulus equal to \(K = 240\) MeV, quite close to the one of the Soft interaction. However the EOS of neutron matter is much stiffer in the BPAL32 case, due to the different behaviour of the potential symmetry energy (see Fig. 1b) in the two interactions considered. In particular, we notice that in BPAL32 the symmetry energy is a monotonic function of the density, in agreement with recent BHF calculations [20,21].

In Fig. 2 we report the solution \(u = \omega/k\) of the d.r. (10) (a) and the corresponding ratio of proton and neutron components (b), \(\delta \rho_p/\delta \rho_n\), for nuclear matter prepared at various density values as a function of the asymmetry \(I\), for the isovector solution. We identify the solution with \(\delta \rho_p/\delta \rho_n > 0\) as the isoscalar solution and the one with \(\delta \rho_p/\delta \rho_n < 0\) as the isovector solution. Here \(\delta \rho_p = \frac{1}{2} (\delta \rho_p + \delta \rho_n), \delta \rho_n = \frac{1}{2} (\delta \rho_p - \delta \rho_n)\).

The solution of the d.r. (10), in the Soft interaction case (left panels), is strongly dependent on the interplay between the Landau parameters, which are functions of density and asymmetry. At densities close to the normal one, the velocity of the isovector mode is an increasing function of the asymmetry and this mode exists up to high asymmetries, as obtained in Refs. [13,22]. But when the density becomes larger than \(\rho_{cross} = 1.4 - 1.5\) \(\rho_0\), where \(F_0(\rho_{cross}) = F_0(\rho_{cross})\) (and the phase velocities of the isovector and isoscalar modes cross each other, c.f. Fig. 3), the isovector mode disappears already at small asymmetries. At densities slightly less than \(\rho_{cross}\), the ratio \(\delta \rho_p/\delta \rho_n\) very quickly tends to 0 as asymmetry increases, that corresponds to mostly neutron oscillations, while at densities larger than \(\rho_{cross}\), this ratio...
becomes less than $-1$, that corresponds mostly to proton oscillations. Indeed, at $\rho_s < \rho_{cross}$, for not too large asymmetries, the phase velocity of the isovector mode is much larger than the Fermi velocity of neutrons and the neutron amplitude dominates because of the neutron excess in matter. Conversely, at $\rho_s > \rho_{cross}$ the velocity of the isovector mode is almost independent on asymmetry, approaching the Fermi velocity of neutrons (see dashed lines in Fig. 2 a,c and Fig. 3 a,c). Thus, we expect here a larger coupling of the collective motion with the single-particle neutron motion. This will induce a gradual suppression of the collectivity, due to the neutron component of the mode. These features are observed also when considering the BPAL32 interaction (right panels). Here $\rho_{cross}$ represents the density value where the phase velocity of the isovector and the isoscalar modes cross each other: $\rho_{cross} = 3.5 \rho_0$.

The only qualitative difference with respect to the case of the Soft interaction is that, with BPAL32 interaction, isovector solutions appear again at high densities $\rho_s > \rho_{cross}$ and finite asymmetries, as a continuation of isoscalar branches (see Fig. 3 c,d).

Using the Soft parameterization, in the region of densities $4\rho_0 < \rho_s < 6\rho_0$ there are no purely real or purely imaginary isovector solutions of the d.r. (10) at $I > 0$ (we did not investigate complex solutions), since the isovector Landau parameter $F'_0$ is very close to zero in this density region. At densities $\rho_s > 6\rho_0$ the isovector solution appears again, but now it is purely imaginary. For densities $6\rho_0 < \rho_s < 10\rho_0$ the isovector mode is overdamped ($\text{Im}(s) < 0$) and at very large densities $\rho_s > 10\rho_0$ it becomes unstable ($\text{Im}(s) > 0$) since the general stability condition:

$$\left(1 + F'^n_0\right)\left(1 + F'^p_0\right) - F'^n_0F'^p_0 > 0$$

(23)

is violated. This condition can be obtained on the basis of thermodynamics arguments [23], but is also follows immediately from the d.r. (11), putting $\omega = 0$. This high density isovector instability could result in the appearance of proton clusters in neutron star matter [1].

On the other hand, according to the interaction BPAL32 (see Fig. 1) the isovector mode will never reach the instability region, due to the linear increase of the potential symmetry energy with density [18].

The isoscalar mode shows the opposite behaviour compared to the isovector one (see Fig. 3). In the low density region the isoscalar mode is unstable, since the condition (23) is broken. At density values $\rho_0 < \rho_s < \rho_{cross}$, the isoscalar mode quickly disappears, diving into the overdamped region. With a further increase of the density, $\rho_s > \rho_{cross}$, the
isoscalar branch becomes robust and exists for all asymmetry values. However, in the case of BPAL32 interaction, this branch experiences, at high density, a gradual transition, with increasing asymmetry, towards solutions of isovector type with $\delta p_x / \delta n < 0$.

In Fig. 4a we present the unstable isoscalar solution of the dispersion relation for low density nuclear matter, as a function of the initial asymmetry. It is possible to observe that the results obtained are almost independent on the introduction of the effective mass in the calculations, as expected in a very dilute system. In Fig. 4b we present the ratio $\delta p_x / \delta n$ as a function of the initial asymmetry. We notice that this ratio is less than $I$. Hence the relative "transition density" of protons is larger than for neutrons, i.e. $\delta p_x / \delta n > \delta p_n / \delta n$, leading to the formation of less asymmetric fragments. This represents a tendency to restore the isotopical symmetry in the spinodal decomposition of neutron rich nuclear matter. Signatures of this effect could be searched by looking at the ratio $N/Z$ of fragments produced in dissipative heavy ion collisions with radioactive beams [24,25].

In conclusion we have shown, in a relatively simple framework, several interesting expectations on charge asymmetry effects in nuclear collective motions. We would like to stress again that, while the appearance of isovector instabilities at high density crucially depends on the interaction used, other important results such as the fragility of isoscalar volume modes in $\beta$-unstable nuclei and the appearance of new unstable oscillations in the low density spinodal region, are quite general features not much depending on the choice of the effective forces.

We are grateful to D.M. Brink and I. Hamamoto for stimulating discussions. One of us (A.B.L.) acknowledges the warm hospitality of the Laboratorio Nazionale del Sud, INFN.

References

|ΔS| = 1 hadronic weak decays of hyperons in a soliton model

Norberto N. Scoccola

Physics Department, Comisión Nacional de Energía Atómica, Av. Libertador 8250, (1429) Buenos Aires, Argentina

Received 22 September 1997; revised 19 March 1998

Abstract

We study the parity violating hyperon non-leptonic weak decays in the three flavor Skyrme model. We follow the approach in which the symmetry breaking terms in the action are diagonalized exactly within the collective coordinate approximation. We show that although this method introduces some configuration mixing, the ΔI = 1/2 rule is numerically well satisfied. In addition, and in contrast to previous calculations, we show that not only the relative amplitudes are in good agreement with the empirical values but also their absolute values. The issue of whether the strong interaction enhancement factors should be included in soliton calculations is also addressed. © 1998 Elsevier Science B.V. All rights reserved.

Nonleptonic weak decays are still one of the least understood aspects of low energy weak interactions. The main difficulty is related with the evaluation of hadronic matrix elements of the weak hamiltonian. In the absence of good hadronic wave functions obtained directly from QCD one has to resort to effective low energy models. In this sense, quark models with QCD enhancement factors have been quite successful in predicting hyperon S-wave decay amplitudes (see Ref. [1] and references therein). The situation in soliton models seemed to be rather different, however. Calculations performed in the mid-eighties showed [2,3] that although octet dominance was present in such models (that is, ΔI = 1/2 rule was well satisfied) and predicted relative amplitudes were in good agreement with the empirical values, their absolute values turned out to be far too small. Such calculations have been done using the so-called “perturbative” approach to the SU(3) Skyrme model. In such an approach, SU(3) collective coordinates are introduced to quantize the soliton and symmetry breaking terms are treated in first order perturbation theory. It is well-known by now that this naive approach leads to very poor predictions even for the hyperon spectra [4]. Moreover, in those calculations the pion decay constant fπ (taken as a free parameter) has to be adjusted to less than one half of its empirical value in order to reproduce some of the observed mass splittings. This small value of fπ was believed to be at the origin of the failure in reproducing the absolute weak decay amplitudes. With the introduction of more refined methods to treat chiral symmetry breaking terms the situation was somewhat improved. In Ref. [5] it was shown that within a framework in which hyperons are treated as soliton-kaon bound systems [6] the calculated matrix elements are indeed larger than those ob-
tained in the perturbative approach. However, they still fall quite below the empirical ones. In this paper we will show that the correct absolute values can be naturally obtained within a scheme in which SU(3) collective coordinates are used but symmetry breaking terms are diagonalized exactly. This approach was pioneered by Yabu and Ando \cite{7} and improved by several authors (for a review see Ref. \cite{8}). As a result of this diagonalization process, configuration mixing appears. One might wonder whether this fact, together with the inclusion of kinematic symmetry breaking terms (needed to obtain good predictions for different observables) will not induce deviations from the empirical well satisfied $\Delta = I = 1/2$ rule. As we will see this is not the case for a reasonable parameterization of the model.

As well-known \cite{9}, using PCAC and isospin symmetry the seven different hyperon non-leptonic amplitudes can be expressed in terms of five independent ones. They are related to the parity conserving weak hamiltonian according to \cite{3}

$$A(A_0) = - \frac{1}{\sqrt{2} f_w} < n|H_{\mu,\Delta S = 1}^p|^A > ,$$

$$A(\Sigma^+) = \frac{1}{2 f_w} < p|H_{\mu,\Delta S = 1}^s|^\Sigma^+ > ,$$

$$A(\Xi^-) = - \frac{1}{\sqrt{2} f_w} < \Lambda|H_{\mu,\Delta S = 1}^s|^\Xi^- > ,$$

$$A(\Sigma^-) = \frac{1}{f_w} < n|H_{\mu,\Delta S = 1}^p|^\Sigma^- > ,$$

$$A(\Sigma^0) = - \frac{1}{\sqrt{2} f_w} [ < p|H_{\mu,\Delta S = 1}^s|^\Sigma^0 >$$

$$+ \sqrt{2} < n|H_{\mu,\Delta S = 1}^p|^\Sigma^0 > ] .$$

For the parity conserving weak interaction hamiltonian we use the Cabibbo current-current form

$$H_{\mu,\Delta S = 1}^p = G \tilde{J}_{\mu,\nu}^L \tilde{J}_{\nu,\mu}^{L^*} .$$

Here, $J_{\mu,\nu}^L$ are the left hadronic currents and $G = G_F \sin \theta \cos \theta \sqrt{2}$, where $G_F$ is the Fermi coupling constant and $\theta$ is the Cabibbo angle. Moreover, we have used the shorthand notation $\pi^- = 1 - i2$ and $\pi^+ = 4 + i5$. Within the Skyrme model the currents $J_{\mu,\nu}^L$ can be obtained as Noether currents of the effective chiral action supplemented with appropriate symmetry breaking terms. We use the form

$$\Gamma = \Gamma_{SK} + \Gamma_{WZ} + \Gamma_{SB} ,$$

where $\Gamma_{SK}$ is the Skyrme action

$$\Gamma_{SK} = \int d^4x \left\{ \frac{f_w^2}{4} \text{Tr} \left[ \bar{\psi} U (\partial^\mu U) \right] \right\}$$

$$+ \frac{1}{32 \epsilon} \text{Tr} \left[ \left( U^{\dagger} \bar{\psi} U, U^{\dagger} \bar{\psi} U \right) \right] .$$

Here, $\epsilon$ is the dimensionless Skyrme parameter. Furthermore the chiral field $U$ is the non–linear realization of the pseudoscalar octet. $\Gamma_{WZ}$ is the Wess-Zumino action:

$$\Gamma_{WZ} = - \frac{i N_c}{240 \pi^2} \int d^4x \epsilon^{\mu\nu\rho\sigma} \text{Tr} \left[ L_\mu L_\nu L_\rho L_\sigma \right]$$

where $L_\mu = U^{\dagger} \bar{\psi} U$ and $N_c = 3$ is the number of colors. Finally, $\Gamma_{SB}$ represents the symmetry breaking terms:

$$\Gamma_{SB} = \int d^4x \left\{ \frac{f^2}{12} \frac{m_a^* + 2 f^2 m_k^2}{m_k^2} \text{Tr} \left[ U + U^\dagger - 2 \right]$$

$$+ \sqrt{3} \frac{f^2}{6} \text{Tr} \left[ \lambda_k \left( U + U^\dagger \right) \right]$$

$$+ \frac{f^2 - f_k^2}{12} \text{Tr} \left[ \left( 1 - \sqrt{3} \lambda_k \right) \right]$$

$$\times \left( U \left( \bar{\psi} U \right)^\dagger \bar{\psi} U + U^{\dagger} \bar{\psi} U \left( \bar{\psi} U \right)^\dagger \right) \right\} .$$

Here $f_k$ is the kaon decay constant while $m_a$ and $m_k$ are the pion and kaon masses, respectively.

A straightforward calculation shows that the corresponding left current can be expressed as

$$J^{L}_{\mu,\nu} = - \frac{i f_w^2}{2} \text{Tr} \left( \lambda_{\mu R} R_\nu \right)$$

$$+ \frac{i}{8 \epsilon^2} \text{Tr} \left[ \left( \lambda_{\mu R} R_\nu \right) \left( R_\rho R_\nu^* \right) \right]$$

$$+ \frac{N_c}{48 \pi^2} \epsilon^{\mu\nu\rho\sigma} \text{Tr} \left( \lambda_{\mu R} R_\rho R_\sigma R_\nu \right)$$

$$- \frac{i f_k^2 - f_w^2}{12} \text{Tr} \left[ \left( 1 - \sqrt{3} \lambda_k \right) \left( U, U \right) \left( R_\mu \right) \right] ,$$

\footnote{Here and in what follows we use the phase convention given in Ref. \cite{1}. Note also that $f_w$ is defined in such a way that empirically $f_w = 93 \text{MeV}$.}
where $R_\mu = \hat{A}_\mu U U^\dagger$. The contribution of the different terms in Eq. (3) to the left current can be easily recognized.

In the soliton picture we are using the strong interaction properties of the low-lying $\frac{3}{2}^+$ and $\frac{5}{2}^-$ baryons are computed following the standard SU(3) collective coordinate approach to the Skyrme model. We introduce the **ansatz**

$$U(r,t) = A(t) \left( \begin{array}{cc} c + i \tau \cdot \hat{F} s & 0 \\ 0 & 1 \end{array} \right) A'(t) \quad (8)$$

for the chiral field. Here we have employed the abbreviations $c = \cos F(r)$ and $s = \sin F(r)$ where $F(r)$ is the chiral angle which parameterizes the soliton. The collective rotation matrix $A(t)$ is SU(3) valued. Substituting the configuration Eq. (8) into $\Gamma$ yields (upon canonical quantization of $A$) the collective Hamiltonian. Its eigenfunctions and eigenvalues are identified as the baryon wavefunctions $\Psi_s(A) = \langle B|A \rangle$ and masses $m_B$. Due the symmetry breaking terms in $\Gamma_{ab}$ this Hamiltonian is obviously not SU(3) symmetric. As shown by Yabu and Ando [7] it can, however, be diagonalized exactly. This diagonalization essentially amounts to admixtures of states from higher dimensional SU(3) representations into the octet ($J = \frac{1}{2}$) and decouplet ($J = \frac{3}{2}$) states. This procedure has proven to be quite successful in describing the hyperon spectrum and static properties [8].

Using the ansatz Eq. (8) in the expression of the left current we obtain that, to leading order in $N_c$, the weak current can be written as

$$H_{\mu A}^{\text{weak}} = -\phi_{3K} R_{\mu A}^{\pi^- S} R_{K^+ S}^{\pi^-} + \phi_{WZ} R_{\pi^- S} R_{K^+ S}^{\pi^-} - \phi_{SB} \left[ \frac{1 + 2 R_{8S}}{2} R_{\pi^- S} R_{K^+ S}^{\pi^-} + \frac{2 + R_{8S}}{3} R_{\pi^- S} R_{K^+ S}^{\pi^-} \right], \quad (9)$$

where

$$\phi_{3K} = \frac{\tilde{G} f_\pi^2}{3} \int d^4 r \left[ \left( \frac{F''^2 + 2 s^2}{r^2} \right) + \frac{4}{e^2 f_\pi^2} \left( \frac{s^2}{r^2} + \frac{2 F''^2 + s^2}{r^2} \right) + \frac{2}{e^2 f_\pi^2} \left( \frac{s^2}{r^2} + \frac{4 F''^2 + s^2}{r^2} + \frac{s^4}{r^4} \right) \right], \quad (10)$$

$$\phi_{WZ} = \frac{\tilde{G} N_c^2}{48 \pi^4} \int d^3 r F' \frac{s^4}{r^4}, \quad (11)$$

$$\phi_{SB} = \frac{\tilde{G} f_\pi^2}{9} \left( f_{K^+}^2 - f_{S}^2 \right) \int d^3 r (1 - c) \times \left[ \left( \frac{F''^2 + 2 s^2}{r^2} \right) + \frac{4}{e^2 f_\pi^2} \left( \frac{s^2}{r^2} + \frac{2 F''^2 + s^2}{r^2} \right) \right]. \quad (12)$$

The SU(3) rotation matrices are defined by

$$R_{a,b} = \frac{1}{2} \text{Tr}(\lambda_a A' \lambda_b A). \quad (13)$$

For simplicity, in Eq. (9) we have not written the contribution quadratic in $(f_{K^+}^2 - f_{S}^2)$ since for empirical values of the decay constants it turns out to be numerically completely negligible.

In the present model the hyperon decay amplitudes can be computed by taking the matrix elements of the hamiltonian Eq. (9) between the hadronic states expressed as linear combinations of SU(3) D-functions. For this purpose it is convenient to use the Clebsch-Gordan decomposition of the collective operators appearing in the weak hamiltonian. One obtains

$$R_{\pi^- S} R_{K^+ S}^{\pi^-} = -\frac{3\sqrt{6}}{5} D_{\frac{3}{2},0}^{\frac{2}{2}} - \frac{1}{10} D_{\frac{1}{2},0}^{\frac{2}{2}} - \frac{1}{\sqrt{20}} D_{\frac{1}{2},0}^{\frac{2}{2}} \quad (14)$$

$$R_{\pi^- S} R_{K^+ S}^{\pi^-} = -\sqrt{\frac{2}{75}} D_{\frac{1}{2},0}^{\frac{2}{2}} - \frac{29}{35} D_{\frac{3}{2},0}^{\frac{2}{2}} - \frac{2}{27 \sqrt{30}} D_{\frac{3}{2},0}^{\frac{2}{2}} + \frac{1}{28 \sqrt{5}} D_{\frac{3}{2},0}^{\frac{2}{2}} - \frac{\sqrt{3}}{21} D_{\frac{5}{2},0}^{\frac{1}{2}}, \quad (15)$$

and similar relations for those containing $R_{\pi^- S} R_{K^+ S}^{\pi^-}$. Here, the left lower index of the SU(3) D-functions $I = \frac{1}{2}, \frac{3}{2}$ stands for $(Y, I, I_1) = (1, I, - \frac{1}{2})$ while the right lower index for $(0, 0, 0)$.

At this stage we note the potential advantages and drawbacks of the present approach with respect to the perturbative calculations of Refs. [2,3]. On one hand the use of an exact diagonalization allows for the use of empirical meson decay constants [8]. This will certainly lead to an improvement of the decay...
amplitudes absolute values. On the other hand, since as a consequence of this diagonalization baryon wavefunctions contain higher SU(3) representations the relevant matrix elements of the collective operators Eqs. (14), (15) will not be, in general, ‘‘octet dominated’’. In this sense, it is not clear whether the $\Delta I = 1/2$ rule will be well satisfied as it was the case in the perturbative calculation.

We turn now to the numerical calculations. We take the meson masses to their empirical values $m_{\pi} = 138$ MeV and $m_{K} = 495$ MeV. Moreover, we use the empirical value $f_\pi = 93$ MeV. As already stressed several times in the literature the use of $f_\pi \neq f_K$ is essential to reproduce the observed mass differences of the low lying octet and decouplet baryons. Therefore, we take $f_K = 120$ MeV which together with $\epsilon = 4.10$ gives a very good overall description of various hyperon properties [8]. As well-known with these parameters the soliton mass turns out to be, at tree level, quite large as compared to the value needed to reproduce the empirical nucleon mass. However, in the last few years it was shown [10] that, within the SU(2) soliton model, the inclusion of one-loop meson corrections reduces that value significantly. Very recently [11] this same conclusion was extended to the SU(3) models. Therefore, at present time, the parameter set above can be considered as the optimal one within the approach adopted here.

Our results for the decay amplitudes are given in Tables 1 and 2. In Table 1 we show the decay amplitude taken with respect to $A(A_0^0)$ while in Table 2 we give the absolute value of this particular amplitude. The results are presented in this way to make easier the comparison with the values obtained in other models. In fact, also shown in Table 1 are those of the perturbative approach (PTA) to the SU(3) soliton model [2], the bound state soliton model (BSA) [5] and the empirical values taken from Ref. [1]. The value for the quark model (QM) that appears in Table 2 has been taken from Ref. [12].

Note that in Table 1 only the values of the independent amplitudes Eq. (2) are given. The reason is that all the corresponding model calculations (and the QM as well) are based on the use of PCAC and isospin symmetry which implies

$$\frac{A(A_0^0)}{A(\Sigma_0^0)} = \frac{A(\Xi_0^0)}{A(\Xi^-)} = \frac{1}{\sqrt{2}}. \quad (16)$$

Although the $A$-ratio is not known empirically, the $\Xi$-ratio is

$$\frac{A(\Xi_0^0)}{A(\Xi^-)}|_{\text{emp}} = -0.75.$$

This is usually taken as an indication that PCAC and isospin symmetry can be used in this framework.

In Table 1 we observe that the relative values of the decay amplitudes are quite well reproduced in our model. Of particular interest is $A(\Sigma_0^0)$. In the limit in which the $\Delta I = 1/2$ rule is exactly satisfied this amplitude is zero. We see that our value, although small, does not vanish. In fact, it nicely reproduces the small departure from the $\Delta I = 1/2$ rule verified by the empirical amplitudes. The reason for the smallness of our calculated value even in the presence of configuration mixing is twofold. Firstly, higher order representations although essential to obtain a reasonable hyperon spectrum appear with a quite small weight in the low-lying hyperon wavefunction. Secondly, the collective operators that contain stronger ‘‘non-octet’’ contributions (as i.e. $R_{sk}R_{a\alpha}^{\sigma'}R_{K^{\alpha'},a}$) appear in terms proportional to $\phi_{sk}$ which is, numerically, one order of magnitude smaller than the leading contributions (terms proportional to $\phi_{sk}$).

### Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$A(\Sigma_0^0)$</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.042 ± 0.006</td>
</tr>
<tr>
<td>$A(\Sigma_0^+)$</td>
<td>-1.26</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00 ± 0.03</td>
</tr>
<tr>
<td>$A(\Sigma^-)$</td>
<td>1.74</td>
<td>1.43</td>
<td>1.41</td>
<td>1.32 ± 0.01</td>
</tr>
<tr>
<td>$A(\Xi^-)$</td>
<td>-1.54</td>
<td>-1.43</td>
<td>-1.73</td>
<td>-1.39 ± 0.01</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>$A(A_0^0) \times 10^6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>This work</td>
<td>0.35</td>
</tr>
<tr>
<td>PTA [2]</td>
<td>0.07</td>
</tr>
<tr>
<td>BSA [5]</td>
<td>0.08</td>
</tr>
<tr>
<td>QM [12]</td>
<td>0.21</td>
</tr>
<tr>
<td>Empirical [16]</td>
<td>0.320 ± 0.002</td>
</tr>
</tbody>
</table>
above, these ‘‘dynamical’’ symmetry breaking terms are important to obtain good mass splittings. Also in Table 1 we observe that the other calculated ratios are (in absolute value) somewhat larger than the empirical ones. However, they are basically of the same quality as those of the PTA or the BSA. In any case, the main success of the present model over the other soliton approaches is in the prediction of the absolute values of the decay amplitudes. Since we have seen that the ratios to the $A(L^0)$ amplitude are reasonably described, it is enough to consider the absolute value of such quantity. From Table 2 we see that our calculated value is in good agreement with the empirical one. The improvement with respect to the PTA and BSA is very significant and shows that the use of empirical values for the model parameters is essential to describe the weak decay amplitudes correctly. In Table 2 we also see that our prediction is somewhat better than that of the QM. However, it should be noticed that in the QM this amplitude is particularly problematic. In general, the QM results are of the same quality than ours.

Finally, we discuss the role of strong interaction enhancement factors used in some previous soliton calculations. These factors were introduced in the context of the quark model to account for hard gluon exchanges [13]. It is not clear whether they should also be used in soliton calculations since, in principle, they could be already contained in the non-perturbative soliton currents. This question was already raised in the context of the $SU(2)$ soliton model (see i.e. Ref. [14]). The results corresponding to PTA and BSA given in Table 2 do include an enhancement factor $c_1 \approx 2.6$. In our calculation we have not used such factor. The satisfactory agreement we have found with respect to the empirical value of $A(L^0)$ suggest that there is no need for the enhancement factors in the soliton models when $H^{pe}_{w,\Delta S=1}$ is defined as in Eq. (2).

In conclusion, we have studied the S-wave nonleptonic weak decay amplitudes of the hyperons in the context of an $SU(3)$ soliton model in which strangeness degrees of freedom are introduced through collective variables and symmetry breaking terms are diagonalized exactly. This approach is known to be successful in describing the hyperon spectrum and other static properties [8]. We have obtained a good agreement of both the relative and absolute amplitudes with the corresponding empirical values. This represents a substantial improvement in the prediction of the absolute amplitudes with respect to previous soliton calculations. Such improvement is greatly due to the use of empirical input parameters. We have also seen that although the present approach includes some configuration mixing the corresponding impact on the ‘‘octet dominance’’ is small enough to guarantee that the empirical ‘‘$\Delta I = 1/2$’’ is still well satisfied. Finally, we note that the simultaneous description of both S-wave and P-wave amplitudes seems to represent a persistent problem in models based in chiral symmetry [15]. Given the satisfactory results found for the parity violating amplitudes it would be interesting to investigate the P-wave decay amplitudes in the present chiral soliton approach.

References

Possible three-nucleon force effects in D–P scattering at low energies

C.R. Brune a,1, W.H. Geist a, H.J. Karwowski a, E.J. Ludwig a, K.D. Veal a, M.H. Wood a, A. Kievsky b, S. Rosati b,c, M. Viviani b

a Department of Physics and Astronomy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3255, USA
and Triangle Universities Nuclear Laboratory, Durham, NC 27708, USA
b Istituto Nazionale di Fisica Nucleare, Piazza Torricelli 2, 56100 Pisa, Italy
c Dipartimento di Fisica, Universitá di Pisa, Piazza Torricelli 2, 56100 Pisa, Italy

Received 13 January 1998; revised 12 March 1998
Editor: R.H. Siemssen

Abstract

We present measurements of the analyzing powers \( A_y \) and \( iT_{11} \) for proton–deuteron scattering at \( E_{c.m.} = 432 \) keV. Calculations using a realistic nucleon-nucleon potential (Argonne V18) are found to underpredict both analyzing powers by \( \approx 40\% \). The inclusion of the Urbana three-nucleon interaction does not significantly modify the calculated analyzing powers. Due to its short range, it is difficult for this three-nucleon interaction to affect \( A_y \) and \( iT_{11} \) at this low energy. The origin of the discrepancy remains an open question.

\[ 0370-2693/98/$19.00 \] © 1998 Published by Elsevier Science B.V. All rights reserved.

PACS: 25.10.+s; 24.70.+s; 21.45.+v

Keywords: N-d scattering; Polarization observables

Few-body systems provide a fundamental testing ground for nuclear interactions. Comparisons of measured three-nucleon scattering observables to theoretical calculations allow stringent tests of the underlying nucleon-nucleon (NN) and models of the three-nucleon (3N) interactions. Past studies have found that rigorous calculations utilizing realistic NN potentials underpredict by 25-30\% the measured analyzing power \( A_y \) in n-d scattering at low energies – a surprising discrepancy which has been dubbed the “\( A_y(\theta) \) puzzle” (see Ref. [1] and references therein). Precise calculations for p-d scattering [2,3] below the deuteron breakup threshold including the Coulomb potential rigorously have recently become possible to perform. These calculations find [3,4] that a similar underprediction exists here for \( A_y \) and also \( iT_{11} \). The agreement for other observables, including cross sections, \( T_{20}, T_{21}, T_{22}, \) and n-d scattering lengths [5] is generally excellent. The analyzing powers \( A_y \) and \( iT_{11} \) in N-d scattering are known to be very sensitive to the NN potential in the \( 1^P \) waves. It has been suggested that the potential in these waves may not be known to the necessary precision at low energies...
The possibility that the underprediction is due to 3N force effects has also been considered [1,3,4]. This paper investigates the roles of NN and 3N force effects on \( A_i \) and \( iT_{11} \) for N-d scattering; we do not consider here other possibly important effects such as relativistic corrections or subnucleonic degrees of freedom. We do note, however, that the Mott-Schwinger interaction, a long-ranged electromagnetic effect, has been recently shown not to be responsible for the discrepancies in \( A_i \) and \( iT_{11} \) [8].

Measurements of both \( A_i \) and \( iT_{11} \) are useful, as these observables depend on different combinations of phase shift and mixing parameters (principally in \( P \)-waves), and they also have been shown to be sensitive to different combinations of the \( ^3P_1 \) NN interactions [6]. The majority of the data on these analyzing powers has been obtained in the vicinity of the deuteron breakup threshold: n-d \( A_i \) data exist for \( E_{cm} \geq 2 \text{ MeV} \) [9,10] p-d \( A_i \) data exist for \( E_{cm} \geq 0.53 \text{ MeV} \) [11-13], and p-d \( iT_{11} \) data exist for \( E_{cm} \geq 1.7 \text{ MeV} \) [12,13] (there are no n-d \( iT_{11} \) data). It is desirable to determine these observables at lower energies, as the influence of higher partial waves is strongly reduced, and the dominant \( S \)- and \( P \)-waves can be investigated with more confidence. In addition, as we will show below, \( P \)-wave N-d scattering at low energies is almost entirely determined by the asymptotic part of the three-nucleon scattering wave function and the NN interaction. Under these conditions the connection between the measured observables and the underlying interactions is greatly simplified. The calculated analyzing powers \( A_i \) and \( iT_{11} \) are mainly determined by the \( j \)-splitting of the \( P \)-wave N-d phase shifts and the \( \varepsilon_{3/2} \) and \( \varepsilon_{1/2} \) mixing parameters. Due to the angular momentum barrier at low energies, these observables are very small and difficult to determine experimentally. This paper reports measurements of \( A_i \) and \( iT_{11} \) for p-d elastic scattering at \( E_{cm} = 432 \text{ keV} \). These data are at the same energy as our previously-reported \( T_{20} \) and \( T_{21} \) measurements [5].

These experimental results are compared to calculations utilizing the Pair-Correlated Hyperspherical Harmonic (PHH) basis [14] to construct the scattering wave function, and the Kohn variational principle to determine the scattering matrix elements [2]. In addition, we present calculations using an “optimized” Born approximation [4] for the peripheral partial waves. The calculations have been done using the AV18 potential [15] and with AV18 plus the 3N interaction of Urbana (UR) [16]. It has been shown in Ref. [1] that other high-quality NN potentials such as Bonn or Nijmegen predict essentially the same n-d \( A_i \) just below the deuteron breakup threshold, so we would not expect our conclusion to change if these potentials were used.

The measurements were performed using polarized proton and deuteron beams from the atomic beam polarized ion source [17] at the Triangle Universities Nuclear Laboratory. The deuteron beams were accelerated to \( E_d = 1.3 \text{ MeV} \) using the FN tandem accelerator, and then directed into a 62-cm diameter scattering chamber. Proton beams were accelerated into a 107-cm diameter scattering chamber at \( E_p = 650 \text{ keV} \) using the mimitandem accelerator [18] and chamber bias voltage [19]. Thin hydrogenated or deuterated carbon targets were utilized which consisted of approximately \( 1 \times 10^{18} \) and \( 1.5 \times 10^{18} \) hydrogen isotope and carbon atoms/cm\(^2\), respectively. The beams lose \( \approx 10 \text{ keV} \) in these targets, with an average energy of \( E_{cm} = 432 \pm 1 \text{ keV} \) for both measurements. The use of thin targets is very important at low energies for minimizing energy loss and straggling effects.

The proton beam polarization was determined using the \( ^6\text{Li}(\vec{p},^4\text{He})^4\text{He} \) reaction in a polarimeter [20] located at the rear of the scattering chamber. The polarization was measured several times throughout the measurements at an incident proton energy of 450 keV by lowering the chamber bias voltage. The proton polarization was found to be constant within \( \pm 3\% \) throughout the measurements; the systematic error in the proton polarization is estimated to be \( \pm 4\% \). Deuteron beam vector polarization was determined online via the \( ^{12}\text{C}(d,p) \) reaction in a polarimeter located behind the scattering chamber. The effective \( iT_{11} \) for this reaction at \( E_d = 1.3 \text{ MeV} \) has been calibrated relative to the \( ^3\text{He}(d,p) \) reaction in another polarimeter at \( E_d = 12 \text{ MeV} \) [21]. The absolute uncertainty in the deuteron beam polarization is estimated to \( \pm 3\% \). For both beams the data were taken with the spin-quantization axis perpendicular to the reaction plane, using two spin states with \( p_z = \pm 0.7 \) for the proton beam; and \( p_z = \pm 0.55 \), \( p_{zz} = 0 \) for the deuteron beam. The spin states were cycled approximately once every
second, in order to minimize the effects of slow changes in beam position, target thickness, or amplifier gain.

Scattered deuterons and protons were detected in coincidence using two pairs of silicon surface barrier detectors placed at symmetric angles on either side of the incident beam. The angles of the detectors were set to observe either protons or deuterons in the more forward detectors in coincidence with deuterons or protons detected in the more backward detector on the opposite side of the beam. Histograms of the time difference between the fast timing signals from each coincident pair of detectors were stored for each spin state. Dead-time corrections (~3%) were determined by sending test pulses to the detector preamplifiers with time delays adjusted to give distinct peaks in the time spectra. The time resolution for the coincident proton-deuteron peaks was ~10 ns, with backgrounds <3%. The analyzing powers were determined from the counts in the coincident peaks, after correction for background, dead time, and the number of incident particles (determined by beam-current integration). It should be noted that the coincidence technique is essential for measuring the small analyzing powers $A$ and $iT_{11}$ with these targets, as the elimination of carbon elastic-scattering events by the fast coincidence requirement allows proton-deuteron scattering events to be counted at the high rate required to achieve reasonable statistical accuracy. The results for $A$ and $iT_{11}$ are shown in Fig. 1. The error bars include contributions from statistics and background subtraction, but not the absolute beam polarization.

The theoretical method has been described previously [2,3]; it can be applied equally well to n-d as well as p-d scattering, and realistic NN and 3N potentials can be used without difficulty. In the present calculations, scattering waves with orbital angular momentum up to $L = 4$ have been taken into consideration. At this energy, the differential cross section, $A$, and $iT_{11}$ are determined almost entirely by waves with $L \leq 1$, while for $T_{20}$, $T_{21}$, and $T_{22}$, $L = 2$ waves are also important. In particular, $A$, and $iT_{11}$ change by $<10^{-4}$ when phases with $L > 1$ are considered. In Fig. 1 the data are compared to the calculations using the AV18 potential and the AV18 + UR potential. The corresponding $P$-wave and $^{3}S_{1/2}$ phase-shift parameters are given in Table 1.

It is seen that both calculations underpredict the data by $\approx 40\%$. The change in the calculated $A$, and $iT_{11}$ resulting from the inclusion of the 3N interaction is too small by an order of magnitude to explain the discrepancy.

Relatively small changes in the N-d phase shift parameters can have large effects on the corresponding analyzing powers. In Ref. [4] it was found that the discrepancies in $A$, and $iT_{11}$ for $E_{\text{cm}} = 1.67$ and 2 MeV could be corrected by reducing the $^{4}P_{1/2}$ phase shift by 3.4% and increasing the absolute value of $\varepsilon_{\frac{1}{2}}$ mixing parameter by 12%. Using the AV18 + UR results for the other phase-shift parameters, we find that agreement with our $A$, and $iT_{11}$ results is optimized if the $^{4}P_{1/2}$ and $\varepsilon_{\frac{1}{2}}$ parameters given in Table 1 (column 2, in parenthesis) are replaced by 5.22 and 1.02, i.e., the absolute values are reduced by 1.6% and increased by 15%, respectively. The results for $A$, and $iT_{11}$ using these parameters are shown by the long-dashed curve in

Fig. 1. Experimental $A$, and $iT_{11}$ for p-d scattering at $E_{\text{cm}} = 432$ keV (circles), along with theoretical calculations using the AV18 (solid line) and the AV18 + UR potentials (short-dashed line). The long-dashed line results from modifying the $^{4}P_{1/2}$ and $\varepsilon_{\frac{1}{2}}$ phases as described in the text.
Table 1

Some of the theoretical phase shifts and mixing parameters (in degrees) calculated at $E_{cm} = 0.432$ and 2 MeV for the AV18 potential; the values in parentheses correspond to the AV18 + UR potential. For both energies, the results for the “optimized” Born approximation and the full solution are reported in the columns labeled “Born” and “Full”, respectively.

<table>
<thead>
<tr>
<th>$E_{cm}$ = 0.432 MeV</th>
<th>$E_{cm}$ = 2 MeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{1/2}$</td>
<td>$s_{1/2}$</td>
</tr>
<tr>
<td>Born</td>
<td>Full</td>
</tr>
<tr>
<td>$-28.9(-28.9)$</td>
<td>$-28.9(-28.2)$</td>
</tr>
<tr>
<td>$-2.15(-2.15)$</td>
<td>$-2.15(-2.15)$</td>
</tr>
<tr>
<td>$-2.01(-2.01)$</td>
<td>$-2.01(-2.01)$</td>
</tr>
<tr>
<td>$4.76(4.76)$</td>
<td>$5.30(5.30)$</td>
</tr>
<tr>
<td>$2\overline{S}_{1/2}$</td>
<td>$2\overline{S}_{1/2}$</td>
</tr>
<tr>
<td>Born</td>
<td>Full</td>
</tr>
<tr>
<td>$-1.64(1.65)$</td>
<td>$2.45(2.50)$</td>
</tr>
<tr>
<td>$-1.99(-1.99)$</td>
<td>$1.99(-1.99)$</td>
</tr>
<tr>
<td>$4\overline{P}_{1/2}$</td>
<td>$4\overline{P}_{1/2}$</td>
</tr>
<tr>
<td>Born</td>
<td>Full</td>
</tr>
<tr>
<td>$5.40(5.39)$</td>
<td>$6.16(6.17)$</td>
</tr>
<tr>
<td>$0.574(-0.573)$</td>
<td>$0.861(-0.888)$</td>
</tr>
</tbody>
</table>

Fig. 1. It is important to note that these parameter changes affect the cross section by < 0.15%, and the other analyzing powers by < 0.0012. In particular, the good agreement observed previously with our $T_{20}$ and $T_{21}$ data at this energy [5] is not disturbed. While similar to the changes required at higher energies [4], there are significant differences in the fractional changes required. We should point out however that there is no reason to expect the percentage change required to be the same for different energies.

We have also performed calculations using an “optimized” Born approximation, i.e., the procedure in which the second-order $\mathcal{M}$-matrix is estimated using the asymptotic part of the three-nucleon wave function as described in Ref. [4]. The results for the $s_{1/2}$ and $P$-wave phase-shift parameters are given in Table 1 and compared to those obtained when the complete wave function is considered (full solution). It is seen that the Born approximation results are close to the full calculation for these partial waves. In the case of the $s_{1/2}$ phase the Pauli principle prevents the three particles from being close to each other, while for the $P$-waves the centrifugal barrier is sufficiently high at these energies. These findings indicate that these partial waves are almost entirely determined by the asymptotic structure of the system. We also show in Table 1 the results for the full solution and Born approximation at $E_{cm} = 2$ MeV, where it is seen that the accuracy of the Born approximation for low partial waves is reduced. On the other hand we observe that the influence of the $\pi$-n force is small and of the same magnitude at both energies.

The small effect on $A_1$ and $iT_{11}$ from including the UR 3N interaction is now clear. This interaction, which is based on two-pion exchange and includes a phenomenological repulsive short range term, requires the three nucleons to be close together in order to produce a significant effect. The likelihood for this situation is diminished by the diffuse structure of the deuteron which results from the small binding energy. For low energy p-d scattering in $P$-waves (or higher $L$ values), the probability of finding three nucleons in close proximity is further reduced by the centrifugal and Coulomb barriers. We thus draw the important conclusion that 3N interactions based on two-pion exchange cannot produce significant changes in $A_1$ and $iT_{11}$ at low energies.

Other choices for the 3N potentials, such as the Tucson-Melbourne [22] or the Brazil [23] models, give quite similar conclusions. Inclusion of other processes, such as $\pi - \rho$ or $\rho - \rho$ exchanges, involving heavier mesons and therefore shorter ranges, are expected to give still smaller corrections [1]. These findings thus indicate that new types of 3N interactions should be considered. One possibility is the inclusion of a spin-orbit 3N force which could significantly affect the N-d $P$-waves [24]. One cannot exclude also the possibility that inadequacies in the NN interaction are responsible for the discrepancy.

In summary, our measurements of $A_1$ and $iT_{11}$ at $E_{cm} = 432$ keV are significantly underpredicted by
calculations utilizing the AV18 NN interaction. The inclusion of the UR 3N interaction does not significantly change the theoretical calculations. We have shown it is difficult to identify a 3N interaction which could significantly change these analyzing powers at low energies, as they are mainly determined by long-ranged interactions. It would be of great interest to extend these comparisons to p-3He and n-3H scattering, where the 3N force effects are expected to be larger, as the likelihood of finding three nucleons close together is enhanced by the tighter binding of 3H and 3He.

Acknowledgements

The authors would like to thank B.M. Fisher for assistance in the data collection process. One of the authors (A.K.) would like to thank Duke University and TUNL for hospitality and partial support during his stay in Durham, where part of the present work was performed. This work was supported in part by the US Department of Energy, Office of High Energy and Nuclear Physics, under grant No. DE-FG02-97ER41041.

References

Pionic double charge exchange on $^{93}$Nb at low energies

J. Pätzold $^a$, R. Bilger $^a$, H. Clement $^a$, A. Denig $^b$, K. Föhl $^{a,2}$, J. Gräter $^a$, W. Kluge $^b$, R. Meier $^a$, M. Schepkin $^c$, G.J. Wagner $^a$, O. Wilhelm $^{a,3}$

$^a$ Physikalisches Institut der Universität Tübingen, Auf der Morgenstelle 14, D-72076 Tübingen, Germany
$^b$ Institut für Experimentelle Kernphysik, Universität Karlsruhe Germany
$^c$ Institute for Theoretical and Experimental Physics, Moscow, Russia

Received 5 February 1998; revised 17 March 1998

Editor: R.H. Siemssen

Abstract

The reaction $^{93}$Nb($\pi^+,\pi^-)^{93}$Te has been measured in the energy range $T_s = 30$–60 MeV at various scattering angles. At all energies the groundstate transition could be observed, whereas the transition to the double isobaric analog state has been identified unambiguously only at $T_s = 50$ MeV. The groundstate transition exhibits a pronounced resonance-like energy dependence which extends the systematics of this peculiar feature to heavy nuclei. An explanation within the $d^\prime$ hypothesis is presented.

1998 Elsevier Science B.V. All rights reserved.

PACS: 25.80.Gn; 27.60.+j; 14.20.Pt

Keywords: Reaction $^{93}$Nb($\pi^+,\pi^-)^{93}$Te; Ground state transition; Transition to double isobaric analog state; Measured $\sigma(E,\Theta)$

The pionic double charge exchange (DCX) reaction at energies below the $\Delta$ resonance has received much attention in recent years. At these energies the cross sections are in general substantially larger than at higher energies and also sensitive to nucleon-nucleon (NN) correlations of short range, a feature which has been looked for since long in this genuine 2N reaction [1]. However, an even more intriguing and completely unexpected observation has been the energy dependence of the forward angle cross sections for monopole transitions in light to medium-heavy nuclei up to $^{56}$Fe. (We note that so far only monopole transitions could be measured at low energies because they are kinematically favoured). Both groundstate transitions (GST) and transitions to double isobaric analog states (DIAT) consistently exhibit a peculiar resonance-shaped structure at energies below the $\Delta$ resonance. The energy of the maximum of this structure shows a systematic dependence on reaction Q-value and the mass number [2]. This phenomenon has not yet found a satisfactory conventional explanation. Even an elaborate coupled-channel treatment of the sequential charge exchange process including explicitly a few individual intermediate nuclear states and carried out for the specific case of the DIAT in $^{14}$C [3] has not been able to provide a.

$^{1}$Supported by the BMBF (06 TU 886), DFG (Mu 705/3, Graduiertenkolleg), NFR and INTAS RFBR (95-605).
$^{2}$Present address: Department of Physics and Astronomy, University of Edinburgh.
$^{3}$Present address: Institut für Hochbautechnik, ETH Zürich.
quantitative description of both energy and angular dependences. Also such a procedure, which crucially depends on details of the structure of individual nuclei, is not likely to be the explanation for a feature common to all nuclei. With a conventional explanation lacking it has been proposed recently [4] that this peculiar phenomenon could be due to the formation of an NN-decoupled $\pi NN$ resonance, called $d'$, with $I(J^P) = \text{even} (0^-), m = 2.06 \, \text{GeV}$ and $T_{\text{NN}} = 0.5 \, \text{MeV}$. In the nuclear medium such a resonance will be broadened by two effects, the center-of-mass motion of the nucleon pair active in the DCX process (“Fermi smearing”) and the collision damping due to $d'N \rightarrow 3N$. Each of these effects gives widths in the order of 10–20 MeV. The $d'$ hypothesis has found further support recently by measurements of $^4\text{He}(\pi^+, \pi^-)pppp$ reaction [5] and in particular by exclusive measurements of the reaction $pp \rightarrow pp\pi^+\pi^-$, where a narrow structure near 2.06 GeV has been observed in the $pp\pi^-$ invariant mass spectrum [6].

In this Letter we report on DCX measurements on $^{93}\text{Nb}$ at various angles and energies. Primary aspect of this experiment has been the extension of the systematics for low-energy DCX into the region of heavy nuclei with emphasis on the question whether the peculiar energy dependence observed in lighter nuclei persists. Hitherto $^{56}\text{Fe}$ has been the heaviest nucleus where angle and energy dependences of the DCX reaction have been studied [7] at energies below the $\Delta$ resonance. We have selected $^{93}\text{Nb}$ as an example of a heavy nucleus for three reasons. First of all, its nuclear structure is fairly simple with three nucleons outside the $Z = 40, N = 50$ closed shell core, the dominant configuration being two neutrons in the $2d_{5/2}$ shell and the single proton in the $1g_{9/2}$ shell. Second, natural Nb is an isotopically pure material making extended targets of purely metallic $^{93}\text{Nb}$ readily available, and third there has been a measurement at LAMPF [8] yielding a surprisingly large cross section of 2.4(3) $\mu$b/str for the DIAT at $T_p = 50$ MeV and $\Theta = 25^\circ$.

Our measurements have been carried out with the Low Energy Pion Spectrometer (LEPS) setup [9] at the $\pi E3$ channel of the Paul Scherrer Institute. As a target we used Nb metal sheets of size $100 \times 100 \, \text{mm}^2$ and areal densities of 428 and 855 mg/cm$^2$, respectively. We note that DCX measurements on heavy nuclei and at low energies are much more difficult than on light nuclei and/or high energies because competing processes may have huge cross sections compared to the DCX process under investigation. In particular, the elastic scattering cross section exceeds that of the GST by a factor of $10^2$ (cf. Figs. 2 and 3 below).

In addition to the magnetic analysis of the negative pions in the LEPS spectrometer, the necessary background reduction has been achieved by time-of-flight measurements of events in the focal plane relative to the LEPS entrance as well as relative to the RF of the cyclotron. Further constraints on particle trajectories and identification are imposed by intermediate and focal plane detectors including a range telescope downstream of the focal plane detector for further $\pi, \mu, e$ separation. Measurements have been performed at two different magnetic field settings of LEPS suitable for the observation of GST and DIAT, respectively. Measurements with the GST setting have been carried out at $T_p = 30–60 \, \text{MeV}$ in the angular range $17^\circ–65^\circ$, measurements with the DIAT setting only at $T_p = 50, 60 \, \text{MeV}$ and $30^\circ$. Sample spectra for both settings are shown in Fig. 1. Whereas the GST peak is observed practically free of background, the DIAT peak sits upon a large

![Fig. 1. Sample momentum spectra for DIAT (left) and GST (right) settings. The left spectrum shows in addition to the DIAT peak an indication of another peak at $Q = -18 \, \text{MeV}$, the expected position of the IAS @ $\text{IAS}$ [8]. The description of the background (breakup channels) is shown by the dashed lines. The solid lines include peaks fitted by Gaussians with experimentally determined widths (s. text).](image-url)
background resulting from breakup channels and unresolved transitions. Also as a consequence of target thickness and the large Q-value of $-22$ MeV, the momentum resolution for the DIAT is somewhat worse than for the GST ($Q = -2.4$ MeV). The widths of the Gaussians fitted to the GST and DIAT peaks have been deduced from elastic scattering measurements at the respective channel energies.

From the evaluation of the spectra absolute cross sections have been obtained by use of the lepton normalization method [9]. As a check for the validity of this method we show in Fig. 2 the simultaneously obtained elastic $\pi^+$ scattering data. The curves represent optical model calculations using the J4 potential [10], which has proven to provide reliable predictions for pion-nucleus scattering at low energies. The data are in full agreement with these predictions. Next we show in Fig. 3 the DCX data obtained for the GST. Measurements with thin and thick targets at $T_p = 50$ MeV and $\Theta = 30^\circ$ agree within statistics and have been combined to one data point. The curves are angular distributions calculated within the $d'$ hypothesis (see discussion below) and normalized in height to the data. This way we deduce forward angle ($\Theta = 5^\circ$) cross sections which allow for a comparison with corresponding LAMPF data [11] taken at higher energies (Fig. 4). As pointed out previously [4], angular distributions calculated with the $d'$ method are compatible in shape with those obtained in conventional DCX calculations which assume the DCX to proceed via sequential SCX processes. In addition, monopole transitions have angular distributions which start with a zero slope at $\Theta = 0^\circ$ and fall off to larger angles as dictated primarily by the size of the target nucleus. Hence the shape of the angular distributions in the forward angle hemisphere can be considered as largely model-independent, so that the extrapolation of our data to $\Theta = 5^\circ$ should be quite safe. The errors
assigned to the \( \sigma(5^0) \) values of Fig. 4 result from a least-square fit of the angular distributions (calculated within the \( d' \) model) to the data (Fig. 3).

The low energy data points of Fig. 4 display the energy dependence of the forward angle cross sections for the GST as obtained from our measurements. For higher energies the LAMPF data [11] are shown. They are consistent with the predictions of the \( \Delta \Delta \) excitation or DINT [1] mechanism as well as with the well-known \( A^{-4/3} \) target mass dependence of GSTs in the \( \Delta \) resonance region. The dashed curve in Fig. 4 shows the result a phenomenological description of this process [12]. Fig. 4 demonstrates rather surprisingly that the new data exhibit a very pronounced structure with a peak cross section at \( T_\pi = 45 \text{ MeV} \) which is an order of magnitude larger than at the \( \Delta \) resonance. In particular there is a fast variation of the cross section by an order of magnitude within only 15 MeV between \( T_\pi = 45 \text{ MeV} \) and 60 MeV.

Before we seek for an explanation of this feature by the \( d' \) hypothesis, we briefly discuss the possibility of other explanations. It has been argued [1,13] that the low-energy structure could be the result of absorption processes increasing very much towards energies in the \( \Delta \) resonance region. However, for heavy nuclei absorption cross sections are fairly constant at low energies. In particular for \(^{93}\text{Nb}\) the absorption cross sections measured [14] at 50 and 85 MeV are of the same size within uncertainties. Hence, it is hard to imagine how a coupling to this channel should create the observed sharp structure in the DCX excitation function. A similar conclusion holds for the possible influence of other inelastic channels. All these channels have cross sections with only gradual changes over the energy region of interest. One may also speculate whether the interesting structure in DCX somehow could be linked to the destructive interference of isovector \( s \) and \( p \) waves, which in light nuclei causes a deep minimum near \( T_\pi \approx 50 \text{ MeV} \) in the forward angle cross section of single charge exchange to isobaric analog states. However, for heavy nuclei this minimum is pretty much washed out and moreover shifted to \( T_\pi \approx 60 \text{ MeV} \) [15].

We turn now to the \( d' \) hypothesis as a possible explanation of the pronounced structure observed in the GST on \(^{93}\text{Nb}\). The solid curve in Fig. 4 represents a calculation with \( m_p = 2.06 \text{ GeV} \), \( \Gamma_{\pi NN} = 0.5 \text{ MeV} \) and \( \Gamma_{\text{spread}} = 15 \text{ MeV} \). The latter accounts for collision damping and fits to the range of values obtained for other nuclei as well as expected from theoretical considerations [2]. The overall phase between \( d' \) and \( \Delta \Delta \) amplitudes has been fixed at 180°. In addition we have increased the \( d' \) amplitude by a factor of two for an optimum fit to the data. This factor should not be considered as a serious problem for the \( d' \) hypothesis, since we have assumed only the simplest shell model transition \( (2 d_{5/2})^2(p_{1/2})^4 \rightarrow (2 d_{5/2})^2(p_{1/2})^3 \) in our calculations for simplicity and it is known that small configuration mixings can already lead to a large enhancement of the \( d' \) amplitude [2,4].

Finally we briefly discuss the experimental situation for the DIAT, which sits upon a large background and hence needs good statistics to be safely identified. For our 60 MeV measurement it turned out that this criterion has not been met. The 50 MeV measurements, however, carried out at \( \Theta = 30^\circ \) with both, thin and thick target exhibit a clear peak at the position of the DIAT \((Q = -22 \text{ MeV}, \text{ see Fig. 1.)} \)

Both measurements give the same cross section within statistical errors resulting in a combined value of 520(90) nb/\( \text{sr}. \) This value, however, is smaller than the corresponding LAMPF result [8] by a factor of five and hence at variance with it. The reason for this huge discrepancy is not known. We have checked the lepton normalization also in the case of these DIAT measurements against elastic pion scattering cross section, which again agree with the J4 predictions. We extrapolate our 30° value to 5° again by use of the \( d' \) model prediction for the angular dependence of the DIAT. This results in \( \sigma(5^0) = 900(150) \) nb/\( \text{sr} \) for the DIAT at \( T_\pi = 50 \text{ MeV} \), a value which is larger than the corresponding cross sections measured in the \( \Delta \) resonance region and above [11], and which fits also reasonably well into the systematics of low-energy DIAT transitions [8,16]. We finally note that one might be inclined to see in Fig. 1, left part, the indication of a further peak structure near \( Q = -18 \text{ MeV}. \) In Ref. [8] it has been argued that this is the transition to the analog of the antianalog state in \(^{93}\text{Mo}\)\((\text{IAS } \otimes \text{ IXS})\). If we take this peak seriously, then its cross section ratio to the DIAT would agree within uncertainties with the one observed in Ref. [8].
In conclusion, the first measurement of the energy dependence of the low-energy DCX on a heavy nucleus exhibits a sharp structure in the forward angle cross section of the GST. Beyond its maximum at $T_s = 45$ MeV the cross section drops by an order of magnitude within $\Delta T_s = 15$ MeV, a feature which appears to be not easily explainable in any known conventional picture. In particular, since all major reaction channels have a smooth energy dependence in this energy region, channel coupling effects appear very unlikely as an explanation. On the other hand this structure fits very well into the systematics found for the low-energy DCX on lighter nuclei and can again be accounted for reasonably well by the $d'$ hypothesis.

References


Magnetic rotation in $^{106}\text{Sn}$ and $^{108}\text{Sn}$

D.G. Jenkins $^a$, I.M. Hibbert $^{a,1}$, C.M. Parry $^a$, R. Wadsworth $^a$, D.B. Fossan $^b$, G.J. Lane $^{b,2}$, J.M. Sears $^b$, J.F. Smith $^b$, R.M. Clark $^c$, R. Krücken $^{c,3}$, I.Y. Lee $^c$, A.O. Macchiavelli $^c$, V.P. Janzen $^d$, J. Cameron $^e$, S. Frauendorf $^f$

$^a$ Department of Physics, University of York, Heslington, York Y01 5DD, UK
$^b$ Department of Physics and Astronomy, SUNY at Stony Brook, Stony Brook, NY 11794, USA
$^c$ Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
$^d$ AECL, Chalk River Laboratories, Ont. K0J 1J0, Canada
$^e$ Department of Physics and Astronomy, McMaster University, Hamilton, Ont. L8S 4M1, Canada
$^f$ FZ Rossendorf, Postfach 510119, D-01314 Dresden, Germany

Received 5 December 1997; revised 6 February 1998
Editor: J.P. Schiffer

Abstract

The nuclei $^{106}\text{Sn}$ and $^{108}\text{Sn}$ have been populated using the $^{54}\text{Fe} (^{38}\text{Ni},\alpha 2p)$ and $^{54}\text{Fe} (^{58}\text{Ni},4p)$ reactions, respectively, at a beam energy of 243 MeV and the gamma rays have been detected using the Gammasphere array. Two “rotation-like” structures consisting of magnetic dipole transitions have been observed in each of the nuclei. The bands can be interpreted, using the Tilted Axis Cranking model, as examples of magnetic rotation. The calculations show excellent agreement with the data for $^{108}\text{Sn}$. However, for the lighter isotope, $^{106}\text{Sn}$, the model is unable to reproduce the experimental $B(M1)/B(E2)$ ratios. Reasons for this discrepancy are discussed. © 1998 Elsevier Science B.V. All rights reserved.

PACS: 21.10.Re; 23.20.Lv; 25.70.Gh; 27.60 +j

The Tilted Axis Cranking (TAC) Model [1] has recently been very successful in describing the properties of regular, “rotation-like” bands, consisting of magnetic dipole transitions, in the light lead isotopes [2–5]. These structures have been interpreted in terms of a new mode of excitation known as magnetic rotation. This term is used since it is the rotation of a large magnetic dipole moment (vector) which breaks the rotational symmetry of the weakly-deformed nucleus, as opposed to the more conventional view of nuclear (electric) rotation where the symmetry is broken by a quadrupole distortion of the electric charge distribution.

The TAC model predicts that magnetic rotation should exist around other proton or neutron closed shell regions [6–9], one of the more promising of
which lies around $Z = 50$ and $A = 110$. In this region the dipole bands might be expected to result from the near perpendicular coupling at the band-head of the angular momentum vectors from a proton configuration involving $g_{9/2}$ holes and neutrons in the $h_{11/2}$ orbitals. The remaining active protons and neutrons would lie in the near degenerate $g_{7/2},d_{3/2}$ shells. The total angular momentum vector, $I$, would then lie at some angle, $\theta$, to the symmetry axis of the weakly deformed nucleus. Higher angular momentum states may be generated by the gradual alignment of the two component vectors with the total angular momentum vector, $I$. The contribution to the total angular momentum from the rotational motion of these near spherical nuclei is expected to be small.

Recent work on $^{105}$Sn [10] has provided the first description in terms of the TAC model of a “rotation-like” band, composed of magnetic dipole transitions, in the tin isotopes. However, the TAC model was only partially successful in interpreting the properties of the band since it failed to reproduce the large experimental $B(M1)/B(E2)$ ratios. In the present work, we have identified four magnetic dipole bands in $^{106,108}$Sn - two bands in each nucleus. The structures in $^{108}$Sn show excellent agreement with TAC calculations. However, as in the case of $^{105}$Sn, only limited success is obtained with the bands in $^{106}$Sn. The results provide important evidence for the phenomenon of magnetic rotation in the light tin isotopes. Reasons for the failure of the model to account for the observed $B(M1)/B(E2)$ ratios in the lightest isotopes, $^{105,106}$Sn, are discussed.

The nuclei $^{105}$Sn and $^{108}$Sn were populated using the $^{54}$Fe($^{58}$Ni, $\alpha$ 2p) and $^{54}$Fe($^{58}$Ni,4p) reactions, respectively, at a beam energy of 243 MeV. The target consisted of 600 $\mu$g/cm$^2$ enriched $^{54}$Fe on a backing of 15.2 $\mu$g/cm$^2$ of gold. This experiment used the Gammasphere array [11] with 95 HPGe detectors. A total of $2.2 \times 10^9$ four-fold or higher events, were collected in prompt coincidence. The data were sorted to produce an $E_x-E_y-E_z$ cube containing $2.1 \times 10^{10}$ triples events, which was analysed using the RADWARE code, Levit8r [12]. The cube was used to extend the previously published level schemes of $^{106,108}$Sn [14–18] and to identify two dipole structures in each of the two nuclei (see Figs. 1 and 2). These data were also unfolded and sorted into three 2D matrices for each nucleus. Matrices were created for all detectors against all detectors, all detectors against those detectors at 90 degrees to the beam axis and all detectors against those at forward (< 80°) angles. The first matrix was used to extract the intensities of the M1 transitions and their respective E2 crossover transitions and, hence, deduce the experimental $B(M1)/B(E2)$ ratios, whilst the latter two matrices were used in order to deduce the spins and parities of states using the method of Directional Correlation of Oriented States (DCO) [13]. DCO ratios were obtained from the ratio of the intensities of transitions in the two matrices, after gating on known stretched quadrupoles on the “all-detector” axis. These values were normalized to take into account the difference between the numbers of detectors at forward angles and those at 90 degrees. The
DCO ratios of known stretched dipole - stretched quadrupole transitions have an average value of 0.85 ± 0.05 whilst known stretched quadrupole - stretched quadrupole transitions have an average value of 1.35 ± 0.05.

In the present work, two structures (labelled as bands 1 and 2) consisting of magnetic dipole transitions and crossover E2 transitions have been identified in $^{108}$Sn (see Figs. 1a and 2a). Dipole transitions from both these structures had been previously observed [14], however, only one (band 1) had been assigned a place in the published level scheme. In the present work we have linked band 2 to the low-lying spherical states and found additional linking transitions for band 1. In addition, we have extended both of these structures and observed E2 crossover transitions in both bands for the first time.

Two structures (labelled as bands 1 and 2) have also been observed in $^{106}$Sn. These consist of magnetic dipole sequences, however, in contrast to the $^{108}$Sn bands, they have no identifiable E2 crossover transitions (see Figs. 1b and 2b). Three of the transitions in band 2 had been observed previously [16] but were unplaced in the level scheme. We have extended this band and determined the links to low-lying states. Band 1 has been observed in several previous experiments [15,16]. In the present work, we have discovered new links for this structure to low-lying spherical states. Furthermore, we identify the band-head of band 1 as the $17^+$ level.
<table>
<thead>
<tr>
<th>Nucleus</th>
<th>$E_r$(keV)</th>
<th>$I_A$(%)</th>
<th>DCO value</th>
<th>Assignment $J^<em>$ → $I_f^</em>$</th>
<th>$B$(M1)/$B$(E2) $(\mu_N/\text{eb})^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>106Sn</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>449.7(1)</td>
<td>30.0(4)</td>
<td>0.87(4)</td>
<td>$18^+ \to 17^+$</td>
<td>&gt; 180</td>
</tr>
<tr>
<td></td>
<td>488.1(1)</td>
<td>17.0(4)</td>
<td>0.80(1)</td>
<td>$19^+ \to 18^+$</td>
<td>&gt; 250</td>
</tr>
<tr>
<td></td>
<td>591.9(1)</td>
<td>9.3(4)</td>
<td>0.75(3)</td>
<td>$20^+ \to 19^+$</td>
<td>&gt; 220</td>
</tr>
<tr>
<td></td>
<td>634.0(3)</td>
<td>6.1(4)</td>
<td>0.69(7)</td>
<td>$21^+ \to 20^+$</td>
<td>&gt; 300</td>
</tr>
<tr>
<td></td>
<td>780.4(5)</td>
<td>1.9(3)</td>
<td>0.84(10)</td>
<td>$22^+ \to 21^+$</td>
<td>&gt; 190</td>
</tr>
<tr>
<td></td>
<td>401.7(2)</td>
<td>6.8(4)</td>
<td>0.84(5)</td>
<td>$18^+ \to 17^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>479.2(2)</td>
<td>17.6(5)</td>
<td>0.91(4)</td>
<td>$19^+ \to 18^+$</td>
<td>&gt; 155</td>
</tr>
<tr>
<td></td>
<td>555.4(2)</td>
<td>7.6(3)</td>
<td>0.77(3)</td>
<td>$20^+ \to 19^+$</td>
<td>&gt; 290</td>
</tr>
<tr>
<td></td>
<td>620.3(2)</td>
<td>4.9(4)</td>
<td>0.83(6)</td>
<td>$21^+ \to 20^+$</td>
<td>&gt; 220</td>
</tr>
<tr>
<td></td>
<td>678.6(4)</td>
<td>1.4(4)</td>
<td>0.96(9)</td>
<td>$22^+ \to 21^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>413.4(1)</td>
<td>27.0(5)</td>
<td>0.70(4)</td>
<td>$15^+ \to 14^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>455.4(1)</td>
<td>53.6(6)</td>
<td>0.81(3)</td>
<td>$14^+ \to 13^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>542.9(1)</td>
<td>36.6(9)</td>
<td>1.14(1)</td>
<td>$17^+ \to 16^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>547.0(1)</td>
<td>31.6(8)</td>
<td>1.10(1)</td>
<td>$16^+ \to 15^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>887.1(1)</td>
<td>17.4(6)</td>
<td>1.43(1)</td>
<td>$14^+ \to 13^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>961.3(1)</td>
<td>70.5(4)</td>
<td>0.86(1)</td>
<td>$13^+ \to 12^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1089.5(2)</td>
<td>2.7(3)</td>
<td>1.40(14)</td>
<td>$15^+ \to 14^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1223.2(4)</td>
<td>1.2(3)</td>
<td>1.3(2)</td>
<td>$17^+ \to 16^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1342.7(2)</td>
<td>2.4(4)</td>
<td>0.18(1)</td>
<td>$14^+ \to 13^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1361.5(2)</td>
<td>3.7(3)</td>
<td>0.92(2)</td>
<td>$13^+ \to 12^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1489.1(2)</td>
<td>1.8(3)</td>
<td>1.20(28)</td>
<td>$17^+ \to 16^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1755.9(2)</td>
<td>2.6(4)</td>
<td>1.50(22)</td>
<td>$15^+ \to 14^+$</td>
<td></td>
</tr>
<tr>
<td><strong>108Sn</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>220.0(1)</td>
<td>1.5(2)</td>
<td>0.83(18)</td>
<td>$13^+ \to 12^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>297.1(1)</td>
<td>4.9(4)</td>
<td>0.80(6)</td>
<td>$14^+ \to 13^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>423.7(1)</td>
<td>10.9(3)</td>
<td>0.95(6)</td>
<td>$15^+ \to 14^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>509.9(1)</td>
<td>9.7(3)</td>
<td>0.86(3)</td>
<td>$16^+ \to 15^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>518.2(2)</td>
<td>5.5(4)</td>
<td>0.93(2)</td>
<td>$17^+ \to 16^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>535.1(1)</td>
<td>1.7(2)</td>
<td>0.75(6)</td>
<td>$18^+ \to 17^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>550.2(2)</td>
<td>1.2(3)</td>
<td>0.75(8)</td>
<td>$19^+ \to 18^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>635.5(4)</td>
<td>0.5(2)</td>
<td>0.91(15)</td>
<td>$20^+ \to 19^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>248.2(5)</td>
<td>0.2(1)</td>
<td>0.95(17)</td>
<td>$17^+ \to 16^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>344.6(4)</td>
<td>0.9(2)</td>
<td>0.76(14)</td>
<td>$18^+ \to 17^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>410.0(3)</td>
<td>3.5(5)</td>
<td>0.92(5)</td>
<td>$19^+ \to 18^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>473.6(3)</td>
<td>2.0(3)</td>
<td>0.83(4)</td>
<td>$20^+ \to 19^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>483.4(5)</td>
<td>2.2(2)</td>
<td>0.90(5)</td>
<td>$21^+ \to 20^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>509.4(8)</td>
<td>1.8(2)</td>
<td>0.72(12)</td>
<td>$22^+ \to 21^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>365.5(5)</td>
<td>1.5(2)</td>
<td>0.76(30)</td>
<td>$18^+ \to 17^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>431.0(1)</td>
<td>7.4(4)</td>
<td>0.62(9)</td>
<td>$14^+ \to 13^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>470.0(8)</td>
<td>1.3(3)</td>
<td>1.85(32)</td>
<td>$19^+ \to 17^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>653.4(1)</td>
<td>4.3(7)</td>
<td>1.4(15)</td>
<td>$14^+ \to 13^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1095.8(2)</td>
<td>11.1(5)</td>
<td>1.48(5)</td>
<td>$15^+ \to 13^+$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1135.6(3)</td>
<td>1.6(3)</td>
<td>1.4(23)</td>
<td>$17^+ \to 15^+$</td>
<td></td>
</tr>
</tbody>
</table>
since the 543 and 547 keV transitions below this state have rather different DCO ratios ($\sim 1.1$) to the transitions above this level (see Table 1). We therefore believe that the transitions below the 17$^-$ level arise from a different structure. DCO ratios for all transitions of interest in both nuclei are presented in Table 1, together with $B(M1)/B(E2)$ ratios for the magnetic dipole bands. The fragmented decay paths of the observed magnetic dipole bands and the DCO ratios of the transitions associated with the decay out of the bands firmly support the spins and parities which have been assigned to these structures.

In the present work, the Tilted Axis Cranking (TAC) model [1] has been used to interpret the structure and properties of the four magnetic dipole bands. Suitable deformation parameters were obtained by running the TAC codes over a series of values for quadrupole deformation, $\varepsilon_2$, and triaxiality parameter, $\gamma$, for a quasiparticle vacuum configuration, in order to minimize the energy in the laboratory frame. For $^{106}$Sn, the minimum corresponded to deformation parameters of $\varepsilon_2 = 0.11$, $\gamma = -13^\circ$, whilst for $^{108}$Sn, the parameters were $\varepsilon_2 = 0.14$, $\gamma = 0^\circ$. The value of the quadrupole-quadrupole coupling constant from Ref. [10] has been adopted in the present work and scaled according to $A^{-2/3}$.

The proton pairing was assumed to be zero due to the proximity of the $Z = 50$ shell gap; accordingly, a single particle model was used to treat the protons. However, the respective neutron numbers $N = 56$ or 58 for the two isotopes implied that neutron pairing effects should be considered. Correspondingly, the pairing constant for neutrons, $\Delta_n$, was set at 1.1 MeV, which is an appropriate value for this region. The chemical potential $\lambda_n$ was set to a value appropriate to each nucleus, i.e. to reproduce either $N \approx 56$ or $N \approx 58$. In order to specify the neutron quasiparticle configuration we use the commonly adopted notation, in which the letters E, F, G and H refer to the lowest $h_\pi$ quasiparticle levels and the letters A, B, C and D refer to the lowest normal parity ($g_{7/2}/d_{5/2}$) quasiparticle levels (e.g. see Fig. 3).

TAC calculations for $^{108}$Sn suggest a structure of $\nu[AE] \otimes \pi[(g_{7/2} \otimes g_{9/2}^{-1})]$ for band 1, whilst band 2 can be explained as resulting from the breaking of a second pair of $g_{7/2}$ neutrons, leading to the configuration $\nu[ABCE] \otimes \pi[(g_{7/2} \otimes g_{9/2}^{-1})]$. This scenario is consistent with the higher excitation energy, larger aligned angular momentum and lower intensity of band 2 relative to band 1. The calculated bandhead excitation energies of 6.0 MeV and 7.5 MeV for bands 1 and 2, respectively, show reasonable agree-
ment with the experimental bandhead energies of 6.665 and 8.102 MeV. Other possible configurations such as $\nu[ACDE] \otimes \pi[(g_{7/2} \otimes g_{9/2})^1]$ have a band excitation energy of over 8.5 MeV.

The calculations for $^{106}$Sn indicate a possible generic configuration appropriate to the two experimentally observed dipole bands, of $\nu[ABCE] \otimes \pi[(g_{7/2} \otimes g_{9/2})^1]$, which generates a solution at a tilting angle, $\theta \sim 23^\circ$ at $\hbar \omega = 0.4$ MeV. Fig. 3 shows that there are two positive parity quasineutron Routhians, labelled B and C, which are very nearly degenerate. It is possible, therefore, to construct variations of the above configuration, for example, configurations based on quasineutrons ABDE and ACDE, which will produce almost identical TAC solutions. Thus, the two bands of the same spin/parity and similar excitation energy, observed in $^{106}$Sn, could be based on such a variation of the generic configuration. All these possible configurations tend to slightly underestimate the increase in spin with rotational frequency when compared to that observed experimentally. However, the calculated excitation energy of the bandhead of 8.7 MeV for the generic $\nu[ABCE] \otimes \pi[(g_{7/2} \otimes g_{9/2})^1]$ configuration, is in reasonable agreement with the experimental bandhead energies of 9.089 and 9.227 MeV for bands 1 and 2, respectively. The bandhead of the $\nu[AE] \otimes \pi[(g_{7/2} \otimes g_{9/2})^1]$ configuration is calculated to lie at about 6.1 MeV. This configuration may be responsible for some of the states in the irregular structure seen below the $17^-$ bandhead of band 1. However, it is more likely that the configuration is unobserved because it is less energetically favourable than the spherical states in this nucleus. This is not an unreasonable scenario given that the deformation of $^{106}$Sn is expected to be smaller than that of $^{108}$Sn. Several other configurations were investigated for both nuclei but the resulting calculations produced much poorer agreement with the experimental data and, moreover, are less likely on the basis of energy considerations.

For $^{105}$Sn, a configuration of $\nu[AEF] \otimes \pi[(g_{7/2} \otimes g_{9/2})^1]$ was previously tentatively assigned to the $\Delta I = 1$ band [10]. However, this assignment is not firm, since the parity of the magnetic rotational band was not firmly established. Furthermore, the quadrupole deformation used in conjunction with the TAC calculations, of $\epsilon_2 = 0.13$, falls between the minimised values for the quadrupole deformation obtained in the present work for $^{106}$Sn ($\epsilon_2 = 0.11$) and $^{108}$Sn ($\epsilon_2 = 0.14$). Calculations for $^{105}$Sn, carried out in the present work, indicate that a quadrupole deformation of $\epsilon_2 = 0.10$, is more appropriate and thereby consistent with the trend seen for $^{106,108}$Sn. Our calculations suggest that the magnetic dipole band in $^{105}$Sn can be satisfactorily explained with a configuration of $\nu[ABE] \otimes \pi[(g_{7/2} \otimes g_{9/2})^1]$ on the alternative assumption that the band is of negative parity. The experimental bandhead excitation energy of 7.043 MeV is consistent with the calculated values for both the present (7.1 MeV) and previous [10] (7.2 MeV) configurations. The assigned configurations, involving one $h_{11/2}$ quasineutron, for the bands in $^{106,108}$Sn are consistent with our alternative assignment for the $^{105}$Sn dipole band. In this scenario, the magnetic dipole bands in all three nuclei may be interpreted as having a $\nu h_{11/2} \otimes \pi[(g_{7/2} \otimes g_{9/2})^1]$ component associated with their configuration; the only difference between them being the number of active $(g_{7/2}, d_{5/2})$ neutron orbitals.

The aligned angular momenta calculated for the proposed configurations in $^{108}$Sn are compared with experimental values in the inset to Fig. 4. Clearly, excellent agreement is found for $^{108}$Sn over a wide rotational frequency range. Both the data and calculations show a nearly linear relationship between $I$ and $E_\gamma$ at low frequencies ($\omega$) which is interrupted by a band-crossing at $\hbar \omega \approx 0.5$ MeV from the alignment of a pair of $h_{11/2}$ neutrons. The agreement between the data and the calculations for $^{106}$Sn is reasonable but not quite as good as for the case of $^{108}$Sn.

TAC calculations for the proposed configurations in $^{108}$Sn predict $B(M1)/B(E2)$ ratios of $\sim 25(\mu_\gamma/\text{eb})^2$ for band 1 and $\sim 15(\mu_\gamma/\text{eb})^2$ for band 2. These values are in good agreement with the experimental values (see Fig. 4). For $^{106}$Sn, the $B(M1)/B(E2)$ values for the configurations discussed, are predicted to be of similar magnitude. Experimentally, however, no $E2$ transitions were clearly distinguishable from the background. The experimental limit obtained for both bands corresponds to $B(M1)/B(E2) \geq 150(\mu_\gamma/\text{eb})^2$ (see Table 1), in disagreement with the calculations. Further calculations were performed in which more positive parity quasineutrons were included in the configura-
Fig. 4. Comparison of calculated and experimental B(M1)/B(E2) values for bands 1 and 2 in $^{108}$Sn as a function of rotational frequency. The configurations used for bands 1 and 2 in $^{108}$Sn are $\pi[AE] \otimes (g_{7/2} \otimes g_{9/2})$ and $\pi[ABCE] \otimes (g_{7/2} \otimes g_{9/2})$, respectively. The inset shows a comparison of calculated and experimental spin against rotational frequency for both bands in $^{108}$Sn.

In conclusion, two magnetic dipole bands have been identified in each of $^{106}$Sn and $^{108}$Sn. The structure of the bands in these nuclei has been interpreted using the Tilted Axis Cranking model.
and an excellent correspondence has been found between theory and experiment for $^{108}\text{Sn}$. However, more limited success was obtained with the bands in $^{106}\text{Sn}$. The principal problems result from the low values of the predicted $B(M1)/B(E2)$ ratios compared to experimental data. Possible reasons for this discrepancy have been discussed in terms of limitations in the way in which the TAC model treats the nuclear deformation and the proton-neutron interaction. These data provide important evidence for the existence of magnetic rotation in the neutron deficient tin region and also highlight some possible deficiencies of the current model.

This work has been supported in part by the U.S. NSF and Department of Energy under Contract No. DE-AC03-76SF00098, the UK EPSRC, AECL and the Canadian NSERC. CMP acknowledges support from the University of York. We would like to thank the staff of the 88–Inch Cyclotron and A. Lipski for making the targets used. The author would also like to thank S. Juutinen of the University of Jyväskylä, Finland, for consultations and helpful discussions during the construction of the level schemes.

References

Warm inflation and classicality conditions

Mauricio Bellini

Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, Funes 3350, (7600) Mar del Plata, Buenos Aires, Argentina

and Instituto de Física y Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo, 58060 Morelia, Michoacán, Mexico

Received 17 December 1997; revised 18 March 1998

Abstract

In this work we analyze the classicality conditions that must be satisfied to have an effective stochastic approach, in the framework of warm inflation scenarios. The radiation temperature at the end of inflation is of the order of $10^{14}$ GeV, for an scalar field with mass $m \sim (10^{-4} - 10^{-5}) M_p$, and the energy fluctuations decreases with the time. © 1998 Published by Elsevier Science B.V. All rights reserved.

PACS: 98.80.Cq; 05.40. + j

1. Introduction

Stochastic inflation has its origin in the mid-80’s, with the papers of Starobinsky and Linde [1]. In this conception of inflation, it was assumed that the universe underwent isentropic expansion during the stage when the scale factor rapidly grew. The entropy required to make the post-inflation universe consistent with observation was assumed to be generated in a short time reheating period. This model predicts that the initial density perturbations should be Gaussian and have a power-law spectrum with index $n \sim 1$. Like other standard slow-roll inflation models, this model separates expansion and reheating into two time periods. Firstly exponential expansion places the universe in a supercooled phase and then the universe is reheated. Recently, A. Berera and Fang [2] have shown how thermal fluctuations during inflation actually may play the dominant role in producing the initial perturbations. They invoked the slow-roll approximation through a combination of a flat potential and dissipative damping. The idea in [2] of introducing a thermal component during inflation was extended in [3] into the warm inflation scenario. In this scenario, slow-roll motion is entirely induced through dissipative damping with no requirement for an ultra-flat potential. In addition, the scenario differs from supercooled inflation scenarios in that reheating is no longer necessary. In the warm inflation scenario, de Sitter expansion and radiation energy production occur together.

I consider a general case, where the inflaton field has a non-null mean value, and develop the analysis by using a consistent semiclassical expansion. As discussed in our previous work [4], this requires
treating the complete second order equation without assuming the slow-roll condition. In addition, following the warm inflation picture, in this work account is made for the interaction of the inflaton with other fields.

The dynamics of a scalar field minimally coupled to a classical gravitational one is described by the Lagrangian:

\[ \mathcal{L}(\varphi, \varphi_{,\mu}) = -\sqrt{-g} \left[ \frac{R}{16\pi} + \frac{1}{2} g^{\mu\nu} \varphi_{,\mu} \varphi_{,\nu} + V(\varphi) \right] + \mathcal{L}_{\text{int}}. \]

In a previous work [4], we have considered how the stochastic approach to inflation relies on the emergence of a long-wave classical field that drives the inflation, and is subject to a shortwave classical noise. In that work we consider explicitly the potential that acts on the inflaton field, and analyze the classicality conditions that must be satisfied to have an effective classical stochastic approach. However, since this is a realization of a supercooled scenario, it is not clear how the quantum to classical transition occurs. The motivation of this work is the study of the quantum to classical transition in the warm inflation scenario. In this context, we add a new term in the Lagrangian (1), that describes the interaction of the field with other fields.

We consider a space-time globally isotropic and homogeneous characterized by a Friedman-Robertson-Walker (FRW) metric, \( ds^2 = -dt^2 + a^2(t)dr^2 \). The resulting equations of motion for the field operator \( \varphi \) and the Hubble parameter \( H \equiv \frac{\dot{a}}{a} \) are:

\[ \ddot{\varphi} - \frac{1}{a^2} \nabla^2 \varphi + (3H + \tau) \dot{\varphi} + V'(\varphi) = 0, \]

\[ H^2 = \frac{8\pi}{3M_p^2} < \dot{\varphi}^2 > \left(1 + \frac{\tau}{4H}\right) + \frac{1}{a^2} < \nabla \varphi >^2 \]

\[ + V(\varphi) > = \frac{8\pi}{3M_p^2} < \rho_r + \rho_\gamma >. \]

where the overdot represents the time derivative, \( V'(\varphi) = dV/d\varphi \) and \( \rho_\gamma, \rho_r \) are the radiation and matter density energy. The term \( \tau(\varphi) \dot{\varphi} \) describes the energy dissipated by the \( \varphi \) field into a thermalized radiation bath.

The expectation value is assumed to be a constant function of the spatial variables for consistency with the FRW metric. From the first law of thermodynamics we have

\[ \dot{\rho}_r + 4H\rho_r = \tau \dot{\varphi}^2. \] (4)

For simplicity, we will consider \( \dot{\rho}_r = 0 \), then

\[ \rho_r = \frac{\tau}{4H} \dot{\varphi}^2. \] (5)

This condition is approximately correct in the warm inflation scenario, and implies that the depletion of the radiation due to expansion will be balanced by its production due to friction. General solutions for \( \rho_r \) and the exact warm inflation scale factor trajectories are computed in [6]. The system we will consider is still dominated by the vacuum energy of the field, with the thermal component being small. Other aspects of our inflation model remain the same as in the standard scenario.

2. The semiclassical expansion of the quantum field

As in a previous work [4], we consider the operator \( \varphi \) as a classical field spatially isotropic (inflaton) \( \phi_i(t) \) plus a quantum fluctuations represented by the operator \( \phi(x,t) \), i.e:

\[ \varphi(x,t) = \phi_i(t) + \phi(x,t), \]

where we require \( <\varphi> = \phi_i(t) \) and \( <\phi(x,t)> = 0 \), for consistency.

2.1. Classical and quantum dynamics

The field \( \phi_i \) is defined as the solution to the classical equation of motion:

\[ \ddot{\phi}_i + (3H_i + \tau_i) \phi_i + V'(\phi_i) = 0, \] (6)

where we have assumed that the classical field is homogeneous, in agreement with the hypothesis of
an inflationary regime. We have interest to estimate the constant $\gamma$ at end of inflation, for
\begin{equation}
\tau_i(0) = \gamma H_o.
\end{equation}
Near the potential’s minimum holds
\begin{equation}
T_s^2 \sim \frac{\tau_i(0)}{4H_o} \phi^2.
\end{equation}
where $\tau_i(0)$ is the value of $\tau_i$ at $\phi_s = 0$.

The evolution of the quantum operator $\phi$ is given by:
\begin{equation}
\ddot{\phi} - \frac{1}{a^2} V\dot{\phi} + (3H + \tau) \phi + \sum_{n} \frac{1}{n!} V^{(n+1)}(\phi_s) \phi^n = 0.
\end{equation}
At the same time the Hubble parameter can be expanded as:
\begin{equation}
H = H_c \left[ 1 + \frac{1}{2H_c^2} \left( 1 + \frac{\tau}{aH_c} \right) \phi^2 + \frac{1}{a^2} (\nabla \phi)^2 + \sum_{n} \frac{1}{n!} V^{(n+1)}(\phi_s) \phi^n \right],
\end{equation}
where:
\begin{equation}
H_c^2 = \frac{4\pi}{3M_p^2} \left[ 1 + \frac{\tau}{aH_c} \right] \phi^2 + 2V(\phi_s)
\end{equation}
is the classical Hubble parameter. If the quantum fluctuations are small the inflation and dissipation are driven by the classical field, and we can restrict the theory to a linear approximation for the quantum fluctuations
\begin{equation}
\ddot{\phi} - \frac{1}{a^2} \nabla^2 \phi + (3H_c + \tau_c) \phi + V''(\phi_s) \phi = 0,
\end{equation}
where we have assumed $H = H_c$ and $\tau = \tau_c$. The classical dynamics of the Hubble parameter being characterized by the equations
\begin{equation}
\dot{\phi}_c = -\frac{M_p^2}{4\pi} \left( H_c + \frac{\tau_c}{3H_c} \right)^{-1}
\end{equation}
\begin{equation}
H_c = \frac{M_p^2}{4\pi} \left( H_c + \frac{\tau_c}{3H_c} \right)^2
\end{equation}
with the potential
\begin{equation}
V(\phi_c) = \frac{3M_p^2}{8\pi} \left[c^2(\tau_c)^2 + \frac{M_p^2}{12\pi} (H_c')^2 \left( 1 + \frac{\tau_c}{4H_c} \right) \right]
\end{equation}
\begin{equation}
\times \left( 1 + \frac{\tau_c}{3H_c} \right)^{-2}.
\end{equation}
Then, the radiation energy density will be
\begin{equation}
\rho_r = \frac{\tau_c}{4H_c} \left( \frac{M_p^2}{4\pi} \right) (H_c')^2 \left( 1 + \frac{\tau_c}{3H_c} \right)^{-2}.
\end{equation}
The density energy fluctuation is given by [2]
\begin{equation}
\rho - \rho_r \sim \frac{\delta \phi V(\phi_s)}{\dot{\phi}^2 + 4/3\rho_r}.
\end{equation}

Now we consider the redefined field $\chi = \exp(3/2((H_c + \tau_c)/dt)\phi)$ that characterize the quantum fluctuations. Then Eq. (12) becomes
\begin{equation}
\ddot{\chi} - \frac{a^{-3/2} \nabla^2 \chi}{a^2 \chi} = 0,
\end{equation}
with
\begin{equation}
k_n^2 = a^2 \left[ \frac{9}{4} \left( H_c + \frac{\tau_c}{3H_c} \right)^2 - V''(\phi_s) + \frac{3}{2} \left( \frac{\dot{H}_c + \dot{\tau}_c}{3H_c} \right) \right].
\end{equation}
The field $\chi$ can be written in terms of the modes $e^{ik \xi_\ell(t)}$,
\begin{equation}
\chi(r,t) = \frac{1}{(2\pi)^{3/2}} \int d^3k \left[ a_k e^{ik \xi_\ell(t)} + h.c. \right],
\end{equation}
where the operators $a_k^\dagger$ and $a_k$ satisfy
\begin{equation}
[a_k^\dagger, a_{k'}] = \delta(k - k'),
\end{equation}
\begin{equation}
[a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0.
\end{equation}
The motion equation for the modes is
\begin{equation}
\ddot{\xi}_\ell + \omega_\ell^2 \xi_\ell = 0.
\end{equation}
Here, $\omega_\ell^2 = a^{-2}(k^2 - k_n^2)$. The modes with $k^2 < k_n^2$
are the responsible of the density inhomogeneities during the inflation.

If \( \chi \) and \( \dot{\chi} \) satisfy the commutation relation
\[
[\chi(r,t),\dot{\chi}(r',t)] = i\delta(r-r'),
\]
then, the time - dependent modes are
\[
\xi_k \xi_k^* - \xi_k^* \xi_k = i,
\]
where the asterisk denotes the complex conjugate.

2.2. The coarse-grained field and classicality

The coarse-grained field
\[
\chi_{cg}(r,t) = \frac{1}{(2\pi)^{3/2}} \int d^3k \theta(ek_o - k) \times \left[ e^{i\dot{\xi}_k(t)} + h.c. \right],
\]

is defined on the long-wavelength modes \((l \geq (ek_o)^{-1}, \varepsilon \ll 1)\). This field satisfy the operatorial stochastic equation \([4]\)
\[
\ddot{\chi}_{cg} - \frac{k^2_o}{a^2} \chi_{cg} = \epsilon \left( \frac{d}{dt}(\dot{k_o}\eta) + 2\dot{k_o}\kappa \right),
\]

with the operatorial noises
\[
\eta(r,t) = \frac{1}{(2\pi)^{3/2}} \int d^3k \delta(ek_o - k) \times \left[ e^{i\dot{\xi}_k(t)} + h.c. \right],
\]
\[
\kappa(r,t) = \frac{1}{(2\pi)^{3/2}} \int d^3k \delta(ek_o - k) \times \left[ e^{i\dot{\xi}_k(t)} + h.c. \right].
\]

When the quantum fluctuations of \( \kappa \) are negligible \( [\kappa \text{ does not commute with the other operators in (26)]} \)
\[
<\dot{k_o}\kappa>^2 \ll <\dot{k_o}\eta>^2,
\]
we obtain the classicality condition for the modes
\[
\left| \frac{\dot{\xi}_k}{\xi_k} \right| \ll 1.
\]

Then, the relation (26) can be written as an semiclassical stochastic equation
\[
\ddot{\chi}_{cg} - \frac{k^2_o}{a^2} \chi_{cg} = \epsilon \frac{d}{dt}(\dot{k_o}\eta).
\]

As in previous work \([4]\), the operators can be redefined such that minimizes the correlation of the noncommuting operator \( \kappa \) \((<\eta \tilde{\eta} + \tilde{\eta} \eta> = 0)\), with
\[
\eta \rightarrow \tilde{\eta} = \left[ 1 + \frac{\dot{k_o}}{k_o} \mathcal{R}_o \right] \eta,
\]
\[
\kappa \rightarrow \tilde{\kappa} = \kappa - \mathcal{R}_o \eta,
\]
where \( \mathcal{R}_o = 1/2 [\dot{\xi}_k/\xi_k + \xi_k^*/\xi_k]_{k=e\kappa} \). The classicality condition with the noises redefined becomes
\[
\left| \xi_{k\eta} \right| < \left| \frac{\dot{k_o}}{k_o} + 2\mathcal{R}_o \right| \ll 1,
\]

and the optimized semiclassical stochastic equation for the coarse-grained field, gives
\[
\ddot{\chi}_{cg} - \frac{k^2_o}{a^2} \chi_{cg} = \epsilon \left[ \dot{k_o}\mathcal{R}_o \eta + \frac{d}{dt}(\dot{k_o}\eta) \right].
\]

Introducing the auxiliar field \( u \), the Eq. (34) can be written as two first order stochastic equations
\[
\ddot{\chi}_{cg} = (\epsilon \dot{k_o}\eta) + u,
\]
\[
\dot{u} = (\epsilon \dot{k_o}\eta) \mathcal{R}_o + \mu^2 \chi_{cg},
\]

where \( \mu^2 = \frac{\lambda}{a} \) is the time dependent mass parameter. The Fokker - Planck equation for the transition probability \( W \) is
\[
\frac{\partial}{\partial t} W = -u \frac{\partial W}{\partial \chi_{cg}} - \mu^2 \chi_{cg} \frac{\partial W}{\partial u} + \frac{\epsilon^2 \dot{k_o}k^2_o}{4\pi^2} \left| \xi_{k\eta} \right|^2 \times \left[ \frac{\partial^2 W}{\partial \chi_{cg}^2} + \frac{2\mathcal{R}_o}{\partial \chi_{cg}} \frac{\partial W}{\partial u} + \frac{\partial^2 W}{\partial u^2} \right].
\]

The following correlation
\[
<\epsilon k_o \eta(t) \epsilon \dot{k_o} \eta(t')> = \frac{e^{i\dot{k_o}k^2_o/2\pi^2} |\xi_{k\eta}|^2}{2\pi^2} \delta(t-t'),
\]
gives the relevant diffusion coefficient:
\[
D^{(\eta\eta)}(t) = \frac{e^{i\dot{k_o}k^2_o/4\pi^2} |\xi_{k\eta}|^2}{4\pi^2}.
\]
3. An example: potential inflation

In this case the scale factor and the Hubble parameter are given by

\[ a(t) = H_0^{-1} t^p, \]  
\[ H_0(t) = \frac{p}{t}. \]

The field's dependence of the Hubble parameter is given by where \( m \) is the scalar field's mass

\[ H_0 = H_0 e^{m \phi}, \]

and the time dependence of the classical field result

\[ \phi(t) = \phi_0 - \frac{1}{m} \ln \left( \frac{H_0}{p} \right). \]

The equation for \( \rho \), is

\[ \frac{M_p^4}{16 \pi^2} \frac{\tau \phi}{H_0} = \rho \left( \frac{\tau}{9 H_0^2} + \frac{2}{3} \frac{1}{H_0} \right) = \rho_\phi. \]

The solution of (44), with for \( \rho_\phi = \frac{1}{8 \pi^2} \left( M_p^2 m H_0 \right)^2 \), result

\[ 1 + \frac{\tau \phi}{3 H_0} = 2 \left( \frac{H_0}{m H_0^2} \right)^2 \left[ 1 - \sqrt{1 - \left( \frac{m H_0^2}{H_0} \right)^2} \right]. \]

When inflation ends, this expression gives

\[ \gamma = \frac{\tau \phi(0)}{H_0} = \frac{3 m^2 H_0^2}{4}, \]

with \( H_0(0) = H_0 = m \). Then, the radiation temperature near the equilibrium is

\[ T_r = \left( \frac{3}{256 \pi^2} \right)^{1/4} (M_p m). \]

that gives \( T_r = 10^{-5} M_p \), for an inflaton with mass \( m = (10^{-4} - 10^{-5}) M_p \). The scalar potential is given by the expression

\[ V(\phi) = \frac{3 M_p^2}{8 \pi} H_0^2 \left( 1 - \frac{M_p^2}{48 \pi} (H_0 m)^2 \right) \left[ \frac{3}{2} \left( \frac{H_0}{m H_0} \right)^2 \right] \times \left[ 1 - \sqrt{1 - \left( \frac{m H_0^2}{H_0} \right)^2} \right] \times \left( 1 - \frac{1}{4} \right) \times \left[ \frac{H_0}{m H_0^2} \right]^{\frac{1}{4}} \times \left[ 1 - \sqrt{1 - \left( \frac{m H_0^2}{H_0} \right)^2} \right]^{\frac{1}{2}}. \]

For \( p \gg 1 \) the equation for the modes can be approximated to with \( p = \frac{4 \pi}{m H_0^2} \)

\[ \ddot{\xi}_k(t) + \left[ \frac{H_0^2 k^2}{\tau^2 p} - \mu^2(t) \right] \dot{\xi}_k(t) = 0, \]

with

\[ \mu^2(t) = \frac{1}{\tau^2} \left( 9 p^2 - 3 \tau (1 + (m)^2) + \frac{13}{8} (m)^3 \right). \]

The normalized solution on the adiabatic vacuum, is

\[ \xi_k(t) = \sqrt{\frac{\tau}{2}} \sqrt{\frac{\pi}{p - 1}} H_0^{(2)} \left[ \frac{H_0}{p - 1} \right]^{\frac{1}{4}} \left( \frac{m H_0^2}{H_0} \right)^{1 - \frac{1}{p}}, \]

where \( H_0^{(2)} \) is the second order Hankel function with

\[ v = \frac{\sqrt{2}}{4(p - 1)} \sqrt{2 + 72 p^2 - 8 \left( m^2 + 3 \right).} \]

Eq. (49) can be written as \( \dot{\xi}_k + \omega_k^2 \dot{\xi}_k = 0 \) [with \( \omega_k^2 = (1/a^2)(k^2 - k_0^2) \)]. Thus, we have a well-defined value of \( k \) \( (k_o) \), which is time dependent threshold between an unstable sector (infrared sector, with \( k^2 < k_0^2 \)) and a stable one \( (k^2 > k_0^2) \). When \( k_o \) surpasses \( k \), the temporal oscillation of the mode ceases. These quantum fluctuations with wave num-
bers below $k_0$ are responsible for the density inhomogeneities generated during the inflation. The condition for the existence of the infrared sector is

$$p \geq \frac{1}{6} + \frac{(m^2)^2}{18} + \frac{\sqrt{3}}{36} \sqrt{18 + 12(m^2) - 115(m^4)}.$$  

The classical fluctuations are given by the solution of the motion equation (where $v = R^{-1}$)

$$\frac{d}{dt} < \phi^2 > = \left[ \frac{\mu^2}{R^2} - 3H_0 \left( 1 + \frac{m^2H_0^2}{4} \right) \right] < \phi^2 > + \frac{\epsilon^2 k_0^2}{4\pi a^3} \left| \xi_{k_{\perp}} \right|^2 + 2R_{c} < \phi v > .$$  \hspace{1cm}(51)

which at the end of inflation (for $t \gg 1$) can be written as

$$\frac{d}{dt} < \phi^2 > = \left[ \frac{\mu^2}{R^2} - 3H_0 \left( 1 + \frac{m^2H_0^2}{4} \right) \right] < \phi^2 > + \frac{\epsilon^2 k_0^2}{4\pi a^3} \left| \xi_{k_{\perp}} \right|^2 ,$$  \hspace{1cm}(52)

with $\mu^2/R^2 = A(p)/t$, such that

$$A(p) = \frac{(1 + 2\nu) \Gamma^2(\nu)}{4\pi} \left( \frac{1}{2} + \nu \right) \mu^2(t) t^{-3(p-1)}.$$

The mapping function is $R(t) = (1/t)$ \times $[1/2 + \nu(p - 1)] + \sigma(\epsilon)$, so that the fluctuations for $p \gg 1$ result (with constant of integration null)

$$< \phi^2 > \sim \frac{\epsilon^2 H_0^2 (p - 1)}{8\pi^2 (A(p) - (3p + 1))} \mu^2(t) t^{-3(p-1)}.$$  \hspace{1cm}(54)

For $p \gg 1$, $< \phi^2 >$ decreases with the time. The energy density fluctuations are

$$\frac{\delta\rho}{\rho} \sim \frac{\epsilon mH_0 (p - 1)^{1/2}}{2\pi (A(p) - (3p + 1))^{1/2}} \mu(t) t^{-3/2(p-1)} ,$$

and the condition for the system to be described by a classical stochastic dynamics is

$$\frac{4\pi (1/2 \epsilon\nu)^{2\nu}}{(1 + 2\nu) \Gamma^2(\nu)} \ll 1,$$  \hspace{1cm}(56)

that satisfies for $\epsilon \ll 1$.

Finally, in this thermal scenario the rapid cooling followed by rapid heating is replaced by a smoothened dissipative mechanism. The warm inflation model studied in this paper predicts energy fluctuations $\frac{\delta\rho}{\rho}$ that decreases with the time for $p \gg 1$.

The evolution of the field fluctuations are dues to

- the cosmological evolution of the horizon and the scale factor [4], and
- dissipation produced by the interaction of the coarse-grained field with the thermal-bath heating of the universe [6] and loss of quantum coherence [5], they are produced by this effect.

When inflation ends, the radiation temperature, $T_r$, is of the order of $10^{14}$ GeV, for a scalar field with mass $m \sim (10^{-4} - 10^{-5}) M_p$.

Acknowledgements

I want to thank the referee for his valuables coments and critisisms, and Rosanna B. Caprioglio for her careful reading of the English of the manuscript.

References


Cosmological perturbation with two scalar fields in reheating after inflation

A. Taruya 1, Y. Nambu 2

Department of Physics, Nagoya University, Chikusa-ku, Nagoya 464-01, Japan

Received 29 August 1997; revised 20 February 1998

Editor: M. Dine

Abstract

We investigate the cosmological perturbation of two-scalar field model during the reheating phase after inflation. Using the exact solution of the perturbation in long-wavelength limit, which is expressed in terms of the background quantities, we analyze the behavior of the metric perturbation. The oscillating inflaton field gives rise to the parametric resonance of the massless scalar field and this leads to the amplification of the iso-curvature mode of the metric perturbations. © 1998 Elsevier Science B.V. All rights reserved.

PACS: 98.80.-k; 98.80.Bp

Keywords: Cosmological perturbation; Reheating; Scalar fields; Parametric amplification

1. Introduction

Recently, a theory of reheating after inflation is developed and importance of the oscillating scalar fields is recognized [1–4]. The non-linear interaction between the scalar fields amplifies the fluctuations of the scalar fields by the effect of parametric resonance and the energy of the oscillating field is transferred to the fluctuation of the massless field by catastrophic particle production. A large amount of studies about the reheating process has been done [5]. From the viewpoint of the structure formation of our universe, one of the most important question is whether the growth of fluctuations of the matter field during reheating affects the large-scale inhomogeneities in the universe or not. To answer this question, we must study the evolution of the metric and the matter fluctuation in general relativistic treatment.

Concerning the works on the theory of cosmological perturbation, several authors studied a single scalar field model in the context of the reheating scenario and investigated a role of the coherent oscillating field on the evolution of large-scale structure in the universe [6–8]. Hamazaki and Kodama analyze the metric perturbation in the model with two-component fluid and discuss the effect of parametric resonance [9]. They evaluate the curvature perturbation ζ (Bardeen parameter) by replacing the scalar fields with the perfect fluids. The effect of parametric resonance is all included in the energy
transfer term which should be given by phenomenological description. They conclude that $\zeta$ is wellconserved during reheating and the parametric resonant decay does not affect the scenario of the structure formation. Although their conclusion seems to be valid in the old version of the reheating scenario, it is not clear that their analysis is appropriate when the non-linear evolution of the scalar fields may play an important role. We should keep in mind that the parametric resonance does not occur in the perfect fluid system.

In this letter, we analyze the cosmological perturbation of the scalar field model to clarify whether the dynamics of coherently oscillating scalar field affects the metric perturbation or not. The model we treat here contains two scalar fields $\phi, \chi$ with the potential $V(\phi, \chi) = m^2 \phi^2/2 + g^2 \phi^2 \chi^2/2$ where $\phi$ is the inflaton and $\chi$ is the massless boson field.

2. Basic equations

We consider a homogeneous, isotropic and flat cosmological model with scalar type metric perturbations. In longitudinal (Newtonian) gauge, the metric is given by

$$ds^2 = -(1 + 2\Phi) dt^2 + a^2(t)(1 - 2\Psi) \delta_{ij} dx^i dx^j.$$  \hspace{1cm} (1)

$\Phi, \Psi$ correspond to the gauge-invariant potential, which characterize the physical mode of the metric perturbations. The Bardeen parameter which represents the spatial curvature perturbation on the uniform energy density slice in super-horizon scale is written by [10–13]

$$\zeta = \Phi - \frac{H^2}{H} (\Phi + H^{-1} \dot{\Phi}) ; \quad H \equiv \frac{\dot{a}}{a}.$$  \hspace{1cm} (2)

For the later analyses, we use the following dimensionless variables:

$$\tau = m t, \quad h = \frac{H}{m}, \quad \varphi_1 = \sqrt{\frac{4\pi}{3}} \frac{\phi}{m_{pl}},$$

$$\varphi_2 = \sqrt{\frac{4\pi}{3}} \frac{\chi}{m_{pl}}, \quad \lambda = \frac{3}{4\pi} \frac{g m_{pl}}{m}.$$  \hspace{1cm} (3)

In terms of these variables, the background Einstein equations are given by

$$h^2 = \varphi_1^2 + \varphi_2^2 + 2U(\varphi_1, \varphi_2);$$

$$U = \frac{1}{2} \varphi_1^2 + \frac{1}{2} \lambda \varphi_1^2 \varphi_2^2.$$  \hspace{1cm} (4)

$$\varphi_1'' + 3h \varphi_1' + \frac{\partial U}{\partial \varphi_1} = 0 \quad (i = 1,2).$$  \hspace{1cm} (5)

where $' = d/d\tau$. We use the following gauge-invariant variable to describe the perturbation:

$$Q_i = \delta \varphi_i - \frac{\varphi_i'}{h} \Psi,$$  \hspace{1cm} (6)

where $\delta \varphi_i$ is perturbation of the scalar field $\varphi_i$. The variable $Q_i$ was first introduced by Mukhanov [14] and all the other gauge-invariant quantities can be expressed by these variables. Using the fact $\Phi = \Psi$ in the present model, the Bardeen parameter defined by (2) is written by

$$\zeta = \frac{h}{\varphi_1^2 + \varphi_2^2} (\varphi_1 Q_1 + \varphi_2 Q_2).$$  \hspace{1cm} (7)

The evolution equation for $Q_i$ is simply expressed by [15]

$$Q_i'' + 3hQ_i' + \left( \frac{k}{ma} \right)^2 Q_i + \sum_{j=1}^2 \left[ \frac{\partial^2 U}{\partial \varphi_i \partial \varphi_j} - \frac{6}{a^3} \left( \frac{a^3}{h} \varphi_i' \varphi_j' \right) \right] Q_j = 0.$$  \hspace{1cm} (8)

This is our basic equation.

3. Solution in long wavelength limit

As we are interested in the formation of the large-scale structure in our universe, it is important to investigate the metric fluctuation outside the scale of the Hubble horizon. Fortunately, we can obtain the exact solution of Eq. (8) in the long wavelength limit ($k \to 0$) and the result is expressed in terms of the background quantities.

Let us suppose the solution of the background field is written as $\varphi_1(\alpha, C)$, where $\alpha = \log a$ and $C$ is the parameter given by a suitable choice of the integration constants. $\alpha$ plays a role of the time
parameter and we can trace the time evolution of the trajectory of the background solution in configuration space \((\varphi_1, \varphi_2)\) using this parameter. \(C\) distinguishes trajectories in this space and remains constant on each trajectory. Using the time parameter \(\alpha\), the field Eq. (5) is rewritten as
\[
\frac{d^2 \varphi_i}{d \alpha^2} + \left( 3 + h^{-1} \frac{dh}{d \alpha} \right) \frac{d \varphi_i}{d \alpha} + h^{-2} \frac{\partial U}{\partial \varphi_i} = 0. \tag{9}
\]
Differentiating (9) with respect to \(\alpha, C\), we can show that the equation of motion for the tangent vectors \(d \varphi_i/d \alpha, \; d \varphi_i/d C\) becomes (8) in the limit \(k \to 0\). Therefore these vectors are the independent solutions of (8) in long wavelength limit [19].

Since (8) are the second order coupled differential equations, we have four independent solutions. The remaining two solutions can be obtained as follows: define the component of a \(2 \times 2\) matrix \(X\) as
\[
X_{11} = \frac{d \varphi_1}{d \alpha}, \quad X_{12} = \frac{d \varphi_1}{d C}, \quad X_{21} = \frac{d \varphi_2}{d \alpha}, \quad X_{22} = \frac{d \varphi_2}{d C}. \tag{10}
\]
Each row vectors in \(X\) satisfy (8). Expressing the remaining solutions by the matrix \(Y\) and putting \(Y = X \cdot P\) yields
\[
P' + (3h + 2X^{-1}X')P' = 0. \tag{11}
\]
We can obtain \(P' = X^{-1}(X^{-1})^T/\alpha^3\) by using the fact that \(X' \cdot X^{-1}\) is the symmetric matrix [19]. Thus the general solution of (8) in long wavelength limit is given by
\[
Q_i = \sum_{j=1}^{2} \left( c_j^{+} X_{ij} + c_j \left[ \int \frac{d \tau}{\alpha} X^{-1}(X^{-1})^T \right]_{ij} \right), \tag{12}
\]
where \(T\) represents transpose of the matrix and \(c_j^{\pm}(j = 1, 2)\) are arbitrary constants.

### 4. Background dynamics in reheating phase

We can know the behavior of the long wavelength perturbation from (12) if we obtain the background solution. As we pay our attention to the reheating phase of the inflaton dominated universe, we assume the condition \(\varphi_1, \varphi_2 \leq 1\) and \(\lambda \varphi_1^2 \leq 1\). The former is the condition for the oscillation of the scalar fields. The latter condition implies that the potential energy is dominated by the massive inflaton.

We first solve the background equations in a naive treatment. Because \(\lambda \varphi_1^2 \leq 1\), the evolution Eqs. (4), (5) can be solved separately and we get oscillatory behavior of the inflaton \(\varphi_1 = \varphi(\tau) \cos \tau\), where \(\varphi \propto 1/\tau\). Substituting this into (5) again, we find that the equation for the field \(\varphi_2\) becomes the same form as the Mathieu equation:
\[
(a^{3/2} \varphi_2)' + [A + 2q \cos(2\tau)](a^{3/2} \varphi_2) = 0; \quad a \propto \tau^{2/3}, \tag{13}
\]
where the coefficients \(A\) and \(q\) are time dependent functions given by \(A = 2q = \lambda \varphi_1^2(\tau)/2\). The stability/instability chart of the Mathieu equation says that \(\varphi_2\) has unstable behavior at \(1/3 \leq q \leq 1\), in the region of the first resonance band of the Mathieu function [19].

To investigate the metric perturbation during the reheating, we must study the background evolution more precisely. We shall analyze the background dynamics by using the renormalization group (RG) method. RG method is a technique of asymptotic analysis which improves the result of the naive perturbation and it provides an unified approach to solve differential equations including singular perturbation method [16, 17].

To apply the RG method, we identify the small expansion parameter \(\epsilon\) as the amplitude of the scalar fields \(\varphi_i \sim \mathcal{O}(\epsilon) < 1\). Since we want to take into account the non-linear interaction, we assume \(\lambda = \epsilon \lambda \sim \mathcal{O}(1)\) which implies \(\lambda\) is not small parameter. Then the background quantities are expanded as follows:
\[
\varphi_i = \epsilon \varphi_{i(1)} + \epsilon^{2} \varphi_{i(2)} + \cdots, \quad h = \epsilon h_{(1)} + \epsilon^{2} h_{(2)} + \cdots. \tag{14}
\]
The crucial treatment in the present model is to insert the mass term \(\omega^2 \varphi_2\) in the equation of \(\varphi_2\) (5). We put \(\omega^2 = 1 + \epsilon \sigma\) and take the limit \(\epsilon \sigma \to -1\) after the end of calculation. This treatment is necessary to extract the effect of parametric resonance in \(\varphi_2\) field.
In appendix, we apply the RG method to the background Eqs. (4), (5). The final form of the solution up to $\mathcal{O}(\epsilon)$ becomes
\[ \varphi_1 = h \cos \Theta \cos(\tau + \psi), \]
\[ \varphi_2 = h \sin \Theta \cos(\tau + \psi), \] (15)
where we replace $\epsilon h_{(1)}$ with $h$. The variables $\Theta, \psi, \phi$ and $h$ are time dependent and determined by so-called RG Eqs. (A.4)–(A.7) which are the result of the renormalization of the initial amplitude.

We further impose the condition $\varphi_2 < \varphi_1$ ($\Theta \ll 1$), which makes our analysis easier. We will restrict our analysis in the case of $\Theta \ll 1$ hereafter. (A.6) becomes $\Theta' = 0$ and we can set $\Theta = 0$ without loss of generality. Eqs. (A.4), (A.5), (A.7) become
\[ \Theta' = \frac{1}{2} q(\tau) \sin \gamma \cdot \Theta, \]
\[ \gamma' = -1 + q(\tau)(2 + \cos \gamma), \] (16)
\[ h' = -\frac{3}{2} h^2, \] (17)
where we define the time dependent parameter
\[ q(\tau) = \frac{1}{4} \lambda h^2. \] (18)

Since (17) yields $h \propto 1/\tau$, $q(\tau)$ is the decreasing function of time. Defining the new variables $u, v$ by $u = \Theta \cos(\gamma/2), v = \Theta \sin(\gamma/2)$, (16) is rewritten as follows:
\[ \begin{pmatrix} u' \\ v' \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & 1 - q \\ 3 q - 1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \] (19)

The eigenvalue of the matrix in the right hand side of (19) is real for $1/3 \leq q \leq 1$ and $\Theta$ can grow during this epoch. We see that the function $q$ corresponds to the coefficient $q$ in the Mathieu equation (13). Recall that we have $\varphi_1 = \Re \cos \tau, \varphi_2 = \Re \Theta \cos(\tau + \gamma/2)$ from (15), we recognize that the amplitude of $\varphi_2$ has the growing behavior as the result of the parametric resonance. Soon after $q$ becomes smaller than $1/3$, $\Theta$ reaches a constant value. Fig. 1 shows the evolution of the background fields obtained by numerical integration of (4), (5). The field $\varphi_2$ (the solid line) is amplified during the short period denoted by the horizontal arrow, which corresponds to the resonance epoch $1/3 \leq q \leq 1$. We can evaluate how much $\varphi_2$ is amplified through the resonance epoch. From (16), we have
\[ \Theta = \Theta_0 \exp \left[ \frac{1}{2} \int_{\tau_i}^{\tau} d\tau' q(\tau') \sin \gamma \right]. \] (20)

Using this expression, we get the upper bound
\[ \frac{\Theta(\tau_f)}{\Theta(\tau_i)} \leq \exp \left( c_s \sqrt{\lambda} \right); c_s = \frac{1}{6} \left( 1 - \frac{1}{\sqrt{3}} \right) = 0.07, \] (21)
where we used the relation $q(\tau_i) = 1/3, q(\tau_f) = 1/10$. The above expression is valid if the condition $\Theta \ll 1$ is satisfied throughout the resonance epoch.

5. Evolution of the Bardeen parameter

Behavior of the background solution obtained by RG method is translated to the metric perturbation $\xi$ using the long wavelength solution (12). The reduced system (16) has the two integration constants: The initial values of $\gamma, \Theta$. During the resonance epoch, $\gamma$ approaches a constant value which does not depend on the initial value. We can regard $\Theta_0$ (the initial
value of $\Theta$) as the parameter $C$ described in Section 3. That is, $\Theta_0$ is the parameter which distinguishes the trajectories in the reduced system ($\Theta, \gamma$). Substituting (15) into (12) and evaluating the integral adiabatically, we find that the dominant contribution to the Bardeen parameter $\zeta$ comes from the solution $Q_{i}^{ad} = d\varphi_i/d\alpha$ and

$$Q_{1}^{iso} = \frac{d\varphi_1}{d\alpha} = -\hbar\Theta_c \sin \theta \cos \tau,$$

$$Q_{2}^{iso} = \frac{d\varphi_2}{d\alpha} = \hbar\Theta_c \cos \theta \cos \left(\tau + \frac{\gamma}{2}\right),$$

where $\Theta_c = \exp\left[\int \frac{d\tau'}{q/2 \cdot \sin \gamma}\right]$. The other independent solutions are recognized as the decaying modes [19].

The solution $Q_{1}^{ad}$ is referred to as the adiabatic growing mode [18]. Substituting $Q_{i}^{ad}$ into (7), the Bardeen parameter remains constant in time. The constancy of $\zeta$ arises in the hydrodynamical perturbation without the entropy fluctuation [12]. For the solution $Q_{i}^{iso}$, (7) gives

$$\zeta^{iso} = \hbar\Theta_c \sin \theta \cos \theta$$

$$\frac{\sin \tau \cos \tau - \sin \left(\tau + \frac{\gamma}{2}\right) \cos \left(\tau + \frac{\gamma}{2}\right)}{\cos^2 \theta \sin^2 \tau + \sin^2 \theta \sin^2 \left(\tau + \frac{\gamma}{2}\right)}.$$  

We see that $\zeta^{iso}$ has the periodically sharp peaks around the zero points of $\varphi_1$. This behavior also appears in the single field case [7]. Except for the short interval of the peaks, $\zeta^{iso}$ traces the amplitude of $Q_{1}^{iso}$. Therefore $\zeta^{iso}$ deviates from zero when $\Theta_c$ grows due to the parametric resonance. Note that $\zeta^{iso}$ vanishes when the background field $\varphi_2$ is zero ($\Theta = 0$). Because the initial amplitude of the curvature perturbation $\zeta^{iso}$ is vanishingly small, we can identify $Q_{1}^{iso}$ as the contribution of the iso-curvature mode of the perturbation.

We confirm the behavior of $\zeta^{iso}$ by solving (4), (5), (8) numerically. In Fig. 2, we show the time evolution of the iso-curvature mode. In contrast to the constancy of the adiabatic growing mode (the dashed line), we observe that $\zeta^{iso}$ (the solid line) is amplified during the resonance epoch when the amplitude of the background $\varphi_2$ field grows (see Fig. 1). After the resonance epoch, $\zeta^{iso}$ approaches zero due to the Hubble damping which can be seen by the analytic result (23). Using (21), (23), we can estimate the growth of $\zeta^{iso}$ except for the sharp peak:

$$\left|\frac{\zeta^{iso}(\tau_f)}{\zeta^{iso}(\tau_i)}\right| \leq \frac{h(\tau_f)}{h(\tau_i)} \cdot \exp\left(2c_s \sqrt{\lambda}\right).$$

6. Summary and discussions

In this letter, we have investigated the cosmological perturbation of the model with two scalar fields in reheating phase after inflation. Using the background quantities, long wavelength solutions of the perturbation are obtained. Applying the RG method to the background dynamics in the coherent oscillating stage, we found that the massless field gets the effect of the parametric resonance. This leads to the amplification of the iso-curvature metric perturbations. The result implies that the Bardeen parameter deviates from the initial amplitude in reheating phase even though the initial amplitude determined in the
inflationary stage is dominated by the adiabatic mode. Therefore the analysis using the scalar field model is essential to study the metric perturbations when the non-linear dynamics of the scalar fields plays an important role.

The most important observed quantity is the power spectrum $\mathcal{P}_k(k)$. Consider the evolution of the metric fluctuation $\zeta(k)$ whose initial amplitude is given by the quantum fluctuation of the scalar fields during inflation. As long as the wavelength of the fluctuation is outside the Hubble horizon, time evolution of $\zeta(k)$ is well-approximated by the solution for the long wavelength mode. The power spectrum for the present length-scale can be evaluated from the amplitude of $\zeta(k)$ when the fluctuations re-enter the Hubble horizon. Soon after the metric fluctuation suffers the parametric amplification during reheating, it approaches to the initial amplitude due to the Hubble damping. The horizon re-entry time becomes earlier as the wavelength of the fluctuation becomes shorter. Therefore, we can expect that the effect of parametric resonance on the power spectrum at the horizon re-entry appears at the small scale (large $k$).

We will give a detail discussion about evolution of the power spectrum in the forthcoming paper [19].

Acknowledgements

This work is partly supported by the Grand-In-Aid for Scientific Research of the Ministry of Education, Science, Sports and Culture of Japan (09740196).

Appendix A. RG method applying to the background system

In this appendix, we apply the RG method to the background system (4), (5). Following the treatment described in Section 4, we substitute (14) to (4), (5) and expand in powers of $e$:

$\mathcal{O}(e^1)$: $\varphi_{i(1)} + \varphi_{i(1)} = 0, \ (i = 1, 2)$

$h_{(1)} = \sqrt{\varphi_{1(1)}^2 + \varphi_{2(1)}^2 + \varphi_{3(1)}^2}$,

$\mathcal{O}(e^2)$: $\varphi_{1(2)} + \varphi_{2(2)} = -3h_{(1)}\varphi_{1(1)} - \lambda\varphi_{2(1)}\varphi_{3(1)}$,

$\varphi_{2(2)} + \varphi_{3(2)} = -3h_{(1)}\varphi_{2(1)} - \sigma\varphi_{2(1)} - \lambda\varphi_{3(1)}\varphi_{2(1)}$.

We first solve the above equations order by order. The solutions up to $\mathcal{O}(e^2)$ become

$$\varphi_1 = e \left[ A_0 + e(\tau - \tau_0) \left\{ -\frac{3}{2}h_{(1)}A_0 
+ i\frac{\lambda}{2} \left( 2|B_0|^2A_0 + B_0^2A_0^* \right) \right] e^{ir}
+ e^2\frac{\lambda}{8}B_0^2A_0 e^{3i\tau} + c.c. + \mathcal{O}(e^3) \right]$$

and

$$\varphi_2 = e \left[ B_0 + e(\tau - \tau_0) \left\{ -\frac{3}{2}h_{(1)}B_0 
+ i\frac{\lambda}{2} \left( 2|A_0|^2B_0 + A_0^2B_0^* \right) \right] e^{ir}
+ e^2\frac{\lambda}{8}A_0^2B_0 e^{3i\tau} + c.c. + \mathcal{O}(e^3) \right],$$

where $c.c.$ means complex conjugate and $h_{(1)}$ is given by

$$h_{(1)} = 2\sqrt{|A_0|^2 + |B_0|^2}.$$  \hspace{1cm} (A.1)

$A_0, B_0$ are the amplitudes at the initial time $\tau = \tau_0$. The above solutions are correct within $e(\tau - \tau_0) \leq 1$ because the secular terms which grow as $(\tau - \tau_0)$ makes the perturbative expansion break down. To improve the perturbation series, we introduce an arbitrary parameter $\mu$. Splitting $\tau - \tau_0$ as $\tau - \mu + \mu - \tau_0$ and absorbing the terms containing $\mu - \tau_0$ into the renormalization constant $Z$, $Z_\theta$ defined by $A_0 = A(\mu)Z_\theta(\tau_0, \mu)$, $B_0 = B(\mu)Z_\theta(\tau_0, \mu)$. This procedure is always possible for each order of expansion and we have the $\mu$ dependent solutions:

$$\varphi_1 = e \left[ A + e(\tau - \mu) \left\{ -\frac{3}{2}h_{(1)}A 
+ i\frac{\lambda}{2} \left( 2|B|^2A + B^2A^* \right) \right] e^{ir}
+ e^2\frac{\lambda}{8}B^2A e^{3i\tau} + c.c. + \mathcal{O}(e^3) \right].$$  \hspace{1cm} (A.2)
\[ \varphi_2 = \epsilon \left[ B + \epsilon (\tau - \mu) \left( -\frac{3}{2} h_{(1)} B + i \frac{\bar{\lambda}}{2} (2|A|^2 B + A^2 B^\ast) + i \frac{\bar{\lambda}}{2} \sigma B \right) \right] e^{i\epsilon t} + \epsilon^2 \frac{\bar{\lambda}}{8} A^2 Be^{i\epsilon t} + c.c. + \mathcal{O}(\epsilon^3). \]  

(A.3)

Since \( \mu \) does not appear in the original equations, the solution should not depend on \( \mu \). This requires \( \partial \varphi_2 / \partial \mu = 0 \) for arbitrary \( \tau \) and we obtain the RG equations:

\[
\frac{dA}{d\mu} = \epsilon \left[ -\frac{3}{2} h_{(1)} A + i \frac{\bar{\lambda}}{2} (2|B|^2 A + B^2 A^\ast) \right] + \mathcal{O}(\epsilon^3),
\]

\[
\frac{dB}{d\mu} = \epsilon \left[ -\frac{3}{2} h_{(1)} B + \frac{i}{2} \sigma B + i \frac{\bar{\lambda}}{2} (2|A|^2 B + A^2 B^\ast) \right] + \mathcal{O}(\epsilon^3).\]

By setting \( \mu = \tau \) in (A.2), (A.3), the secular terms are eliminated and we obtain the regular perturbation. Rewriting the variables \( A, B \) by

\[ A = \frac{1}{2} h_{(1)} \cos \Theta e^{i\phi}, \quad B = \frac{1}{2} h_{(1)} \sin \Theta e^{i\phi}, \]

and using (A.1), we get the solution (15) and the above RG equations reduce to

\[ \Theta' = \epsilon \frac{\bar{\lambda}}{16} h_{(1)}^2 \sin \gamma \sin 2\Theta, \quad \gamma = 2(\psi - \theta), \]

\[ \gamma' = -1 + \epsilon \frac{\bar{\lambda}^2}{4} h_{(1)}^2 \cos 2\Theta (2 + \cos \gamma), \]

(A.4)

\[ \theta' = \epsilon \frac{\bar{\lambda}}{8} h_{(1)}^2 \cos \gamma \sin^2 \Theta, \]

(A.6)

\[ \gamma' = -1 + \epsilon \frac{\bar{\lambda}^2}{2} h_{(1)}^2, \]

(A.7)

where we set \( \epsilon \sigma = -1 \).

References

A note on inflationary string cosmology

Stefan Förste

Sektion Physik, Universität München, Theresienstraße 37, 80333 München, Germany

Received 9 March 1998
Editor: L. Alvarez-Gaumé

Abstract

Cosmological solutions are obtained by continuation of black D-brane solutions into the region between the horizons. It is investigated whether one can find exponential expansion when probing the cosmology with D-branes. A unique configuration exhibiting exponential expansion is discussed. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

String (or M-) theory is believed to describe physics at the Planck scale. The only experimental window into this region is the early universe. Therefore it is very interesting to derive implications for the evolution of the universe from string theory. In [1–16] we give a list of some references dealing with this question. A prominent stringy version of a cosmological model is the pre-big-bang (PBB) scenario [2] ². There, two solutions connected by scale factor duality are used to describe an inflationary and a Friedmann-Robertson-Walker (FRW) phase of the universe. One of the main problems in the PBB scenario is to describe the smooth phase transition from the inflationary to the FRW phase. This is known as the graceful exit problem [3]. A related problem is that typically cosmological solutions of low energy effective string theories run into singularities. Methods of how to avoid those singularities are for example discussed in [4–6,10]. Furthermore, there are recent indications that the PBB scenario also requires an exponentially large world radius at its onset [16].

In this note we are going to investigate whether we can find solutions exhibiting an inflationary phase with exponentially growing world radius. We will deal with solutions obtained in the spirit of [7], i.e. by continuing a black D-brane solution into the region between inner and outer horizon. Black D-brane solutions minimize the type II low energy effective action,

\[ S = \int d^{10} x \sqrt{-g} \left[ e^{-2\phi} \left[ R + 4(\partial \phi)^2 \right] - \frac{2}{(8 - p)!} F^2 \right]. \]

(1)
$F$ is a RR-form field strength. An object extended along $p$ spatial dimensions couples electrically to a $p+1$ form gauge potential, and thus corresponds to non-zero $p+2$ form field strength. Here, we will discuss magnetically charged objects and hence will have non-zero $8-p$ form field strength. In [13] it has been observed that for an $(n,m)$ five-brane of type IIB theory the presence of RR charges increases the acceleration of the expansion near the big bang singularity. Therefore, there is some hope that one can find exponential expansion from D-brane backgrounds. In section two we will scan all the black D-brane backgrounds (continued to the region between the horizons) for exponentially growing world radii near the singularity at the inner horizon. The scan includes T-dual solutions (or, in other words D-branes whose transverse space has compact directions), and the possibility of probing the universe with several kinds of D-branes. We will find that there is exactly one option to obtain exponential inflation within the given framework. This is a T-dualized D5-brane solution probed by D0-branes. (After T-duality the D5-brane becomes a D4-brane and therefore probing with D0-branes makes sense.) In section three we will investigate the inflationary model and find that it expands in all spatial directions and is therefore a truly ten dimensional solution. We comment on our results in a concluding section.

2. Scan for inflationary string cosmology

The strategy is as follows. We take a black $p$-brane solution and ‘go’ in the region between the horizons. In the region between the inner and the outer horizon, time becomes space-like and the radial distance becomes time-like. Thus the solution (depending on the radial distance) will become a time dependent cosmological solution. Typically, the solutions obtained in this way possess a big bang singularity when time is at the inner horizon and a big crunch singularity when it reaches the outer horizon (or vice versa). Our aim is to investigate whether we can find frames with an exponential expansion in the vicinity of the inner horizon. In such frames the universe would exhibit inflationary expansion close to the big bang singularity. Especially we would like to find solutions for which three spatial directions exponentially expand whereas the others contract or at least do not expand.

Before discussing the general black D-brane solution we describe what we mean by brane frame. A brane frame is defined such that the world volume action of the corresponding D-brane starts off with the ‘canonical’ Nambu-Goto term,

$$\int d^{p+1}x \sqrt{g_{\text{brane}}} \equiv \int d^{p+1}x e^{-\phi_0} \sqrt{g_{\text{string}}}$$

where in our convention $e^{\phi_0}$ is the string coupling and the index $i$ on the metrics refers to the fact that they are induced. From the above we obtain for the metric components

$$(-2 \phi_0 + 2 \phi_i)$$

When we probe space-time with a D-p-brane we will measure the corresponding brane-frame metric.

Now we will scan systematically all D-q-brane backgrounds for the possibility of inflationary phases in some p-brane frame. For that we use the general form of the black brane solution in [17]. From their solution we
obtain a cosmological solution by continuation into the region between the two horizons. We change the notation of [17] according to

\[ t \rightarrow y, \quad r \rightarrow t, \quad r_{\pm} \rightarrow t_{\pm}, \quad (t_{+} > t_{-}) \]

and chose \( t \) to be in the interval between \( t_{-} \) and \( t_{+} \). For a D-q-brane we find for the metric in the string frame

\[ ds^2 = \left( \left( \frac{t_{+}}{t} \right)^{7-q} - 1 \right) \sqrt{\frac{t_{-}}{(7-q) e}} \, dy^2 - \frac{dt^2}{\sqrt{\frac{t_{-}}{(7-q) e}} \left( \left( \frac{t_{+}}{t} \right)^{7-q} - 1 \right)} + \frac{1}{2} \frac{5-q}{7-q} d\Omega_{8-q}^2 + \frac{1}{2} \frac{5-q}{7-q} e \, dx^i \, dx^i, \]

where the sum over \( i = 1, \ldots, q \) is understood. The dilaton is (for convenience we put \( \phi_0 = 0 \), it can be reintroduced by noting that constant shifts in \( \phi \) are moduli of the low energy effective theory)

\[ e^{-2\phi} = \left( 1 - \left( \frac{t_{+}}{t} \right)^{7-q} \right)^{\frac{3-q}{2}}. \]

For a magnetically charged q-brane one has an \( 8-q \) form field strength,

\[ F = Q e_{8-q} \]

where \( e_{8-q} \) is the unit volume form of a \( 8-q \) sphere \(^4\) and the value for \( Q \) is

\[ Q = \frac{1}{2} \left( \frac{7-q}{2} \right)^{2} \left( t_{+}, t_{-} \right)^{7-q}. \]

Now we consider the region near the inner horizon, \( t = t_{-} + \epsilon \),

\[ 1 - \left( \frac{t_{+}}{t} \right)^{7-q} = (7-q) e_{t_{-}} + \mathcal{O}(\epsilon^2). \]

In this region the metric is approximated by

\[ ds^2 = \left( \left( \frac{t_{+}}{t_{-}} \right)^{7-q} - 1 \right) \sqrt{\frac{t_{-}}{(7-q) e}} \, dy^2 - \frac{dt^2}{\sqrt{\frac{t_{-}}{(7-q) e}} \left( \left( \frac{t_{+}}{t_{-}} \right)^{7-q} - 1 \right)} + \frac{1}{2} \frac{5-q}{7-q} d\Omega_{8-q}^2 + \frac{1}{2} \frac{5-q}{7-q} e \, dx^i \, dx^i. \]

After compactifying the \( x^i \) and \( y \) directions (spanning the Euclidean world-volume of the brane) on circles, T-duality can be performed along the \( x^i \) directions or along the \( y \) direction. The \( g_{tt} \) component of the metric is not affected by these T-dualities and will always behave like

\[ g_{tt} \sim e^{\frac{1}{2} \frac{5-q}{7-q}}. \]

\(^4\) For \( q = 3 \) one replaces \( F \) by \( F + * F \) in order to have a self-dual field strength [17].
The other metric components show generically some power-like behavior with respect to $e$. In order to get exponential expansion (or contraction) in a proper time frame the time component of the metric should go like $e^{-2}$. To find inflation for a $p$-brane probe we need thus a dilaton behaving like (the tilde on $\phi$ indicates that we allow for T-duality transformations)

$$e^{-2\tilde{\phi}} \sim e^{(-\frac{1}{2} + \frac{5 - q}{7 - q}(p + 1))},$$

where $p + 1$ is the world-volume dimension of the probe. Using the T-duality relation

$$\tilde{\phi} = \phi + \frac{1}{2} \log \frac{\tilde{g}}{g}$$

we observe that performing T-duality with respect to $n$ of the $x^i$ ($n = 0, \ldots, q$) and $m$ times with respect to $y$ ($m = 0, 1$) will lead to the following expression:

$$e^{-2\tilde{\phi}} \sim e^{-\frac{3 - q}{2} - \frac{n + m}{2}}.$$ (12)

Comparing (12) with (14) we arrive at the condition

$$\frac{3 - q}{2} + \frac{n}{2} - \frac{m}{2} = (p + 1)\left(-\frac{1}{2} + \frac{5 - q}{7 - q}\right).$$ (15)

Within the allowed parameter regions there is only one solution to (15), namely $q = 5$, $p = 0$, $m = 1$, $n = 0$. (16)

So, we have to start with the D5-brane background. By T-dualizing the $y$ direction this will turn into a D4-brane (with world-volume along $x^i$, $i = 1, \ldots, 5$). Probing this D4-brane background with D0-branes will result in measuring exponentially growing scale factors near the singularity at the inner horizon. (Later we will see that in order to get expansion and not contraction we have to reverse the time direction. Then the singularity at the inner horizon corresponds to a final singularity which will occur in the infinite future of a proper time.) It is quite surprising that we find exactly one possibility to obtain inflation within the given framework. Fortunately, the background and the probes are both objects of type IIA theory.

3. Inflationary string cosmology from the D5-brane

In the previous section we gave general arguments that starting with a (continued) D5-brane solution, performing T-duality along the $y$ direction and moving to the D0-brane frame will give inflationary cosmology near the singularity at $t = t_*$. Here, we will repeat the previous discussion for the D5-brane and analyze the result. We start with the solution

$$ds^2 = \frac{\left(\left(\frac{t}{t_*}\right)^2 - 1\right)}{\sqrt{1 - \left(\frac{t}{t_*}\right)^2}} dy^2 - \frac{dt^2}{\left(\left(\frac{t}{t_*}\right)^2 - 1\right)^{\frac{1}{2}}} + t^2 \left(1 - \left(\frac{t}{t_*}\right)^2\right)^{-\frac{1}{2}} d\Omega_3^2 + \sqrt{1 - \left(\frac{t}{t_*}\right)^2} dx^i dx^i,$$ (17)

Formally (15) is also solved by $q = 8$, $n = 8$, $m = p = 0$. The D8-brane is a solution of massive type IIA supergravity [18] and not of the effective theories considered here.
The three-form RR-field strength is
\[ F = Q \epsilon_3 \]
where \( \epsilon_3 \) is the unit volume form of a three-sphere and
\[ Q = t_, t_. \]
T-dualizing along the \( y \) direction yields
\[ ds^2 = \sqrt{1 - \left( \frac{t_-}{t} \right)^2} dy^2 - \frac{dt^2}{\left( \left( \frac{t_+}{t} \right)^2 - 1 \right)} + t^2 \left( 1 - \left( \frac{t_-}{t} \right)^2 \right)^{\frac{1}{2}} d\Omega_3^2 \]
\[ + \sqrt{1 - \left( \frac{t_-}{t} \right)^2} dx'dx', \]
and a four form field strength (whose gauge field couples magnetically to the D4-brane),
\[ F \sim \epsilon_3 \wedge dy. \]
Finally we move to the D0-brane frame by replacing \( ds^2 \rightarrow ds_0^2 = e^{-2\phi} ds^2 \),
\[ ds_0^2 = \frac{1}{1 - \left( \frac{t_-}{t} \right)^2} dy^2 - \frac{dt^2}{\left( 1 - \left( \frac{t_-}{t} \right)^2 \right)^{\frac{2}{3}}} + t^2 \left( \left( \frac{t_+}{t} \right)^2 - 1 \right) d\Omega_3^2 + \frac{\left( t_+ \right)^2 - 1}{1 - \left( \frac{t_-}{t} \right)^2} dx'dx'. \]
In order to discuss the characteristics of (24) we transform to the proper time coordinate defined via
\[ d\tau = \pm \frac{dt}{1 - \left( \frac{t_-}{t} \right)^2}, \]
which is solved by
\[ \pm \tau = t + \frac{t_-}{2} \log \left( \frac{t - t_-}{t + t_-} \right) + \text{constant}. \]
Unfortunately we cannot obtain an analytic expression for \( t \) as a function of \( \tau \) and therefore we will just give a qualitative discussion. We reverse the time direction by choosing the lower sign in (26) and take the initial condition such that
\[ t(\tau_0) = t_. \]
Now the big bang is at \( t = t_+ \) where the scale factor in front of \( dx'dx' \) and \( d\Omega_3^2 \) vanishes. Near the initial
value $\tau_0$, $t$ depends linearly on $\tau$. As we approach $t_-$ all the scale factors grow exponentially with $\tau$ and diverge at $t(\tau = \infty) = t_-$. To illustrate this we give the approximate metric near $t = t_-$. In that region we can solve for $t(\tau)$ by

$$t(\tau) \approx t_- + c t_- e^{-\frac{2\tau}{t_-}},$$

with $c$ being some integration constant depending on the initial condition (27). Then the metric becomes

$$ds_5^2 = -d\tau^2 + \frac{e^{\frac{\tau}{t_-}}}{2c} \left( dy^2 + \left( \frac{t_+}{t_-} \right)^2 - 1 \right) \left[ t_+^2 d\Omega_3^2 + dx'dx'' \right].$$

(29)

The dilaton behaves like

$$e^{-2\phi} \approx \left( \frac{t_+}{t_-} \right)^2 \left( 2c \right)^{\frac{3}{2}} e^{\frac{2\tau}{t_-}},$$

(30)

and large $\tau$ is seen to correspond to weak coupling. However, for large $\tau$ the curvature increases and $\alpha'$ corrections will become important. (Note that we are dealing with non-BPS states.) To summarize, we have found a cosmological solution which after some power-like expansion enters an inflationary phase with exponential expansion, when probed with D0-branes. For large proper time the curvature will blow up and $\alpha'$ corrections will be important. In the presented solution all the nine spatial directions become exponentially large.

4. Conclusions

In this note we used the fact that a black D-brane corresponds to a cosmological solution when continued to the region between the inner and the outer horizon. From these cosmological solutions new solutions can be generated via T-duality. Then we searched this class of cosmological solutions for a metric describing exponential expansion when probed with D-branes. Surprisingly we found exactly one solution. This is a D5-brane T-dualized to a D4-brane. When this background is probed with D0-branes, the D0-brane like observer will see a universe which enters an exponential expansion after some time. Since all nine space directions expand exponentially the universe is truly ten dimensional.

One may be tempted to use the presented solution in some kind of a double-pre big bang scenario in order to achieve an exponentially large world radius at the onset of the usual pre-big-bang scenario. However, first of all there will be problems because one has to introduce a second graceful exit from the exponential expansion to the inflationary expansion of the PBB scenario. This may be merely a technical question. But even if one was able to solve for the graceful exit one would have just reversed the problem raised in [16]. Now, we would need an exponential contraction in order to get the non-observable six space dimensions down to string scale. Before interpreting this result as a no-go theorem for phenomenologically interesting string cosmology one should note that we only considered a special class of cosmological string vacua. Extending the scan for exponential expansion to more general solutions one might be more lucky in finding some model of phenomenological relevance. As a step towards that direction one may consider for example intersecting black branes [19] as a starting point.

Acknowledgements

It is a pleasure to thank Debashis Ghoshal, Jacek Pawelczyk and Stefan Theisen for many valuable discussions. I also thank Stefan Theisen for a critical reading of the manuscript. The author is supported by GIF
– the German Israeli Foundation for Scientific Research. The work is also supported in part by TMR program ERBFMX-CT96-0045.

References


Planckian scattering of D-branes

Saurya Das a,b,1, Arundhati Dasgupta a,2, P. Ramadevi c,3, Tapobrata Sarkar a,4

a The Institute of Mathematical Sciences, CIT Campus, Chennai - 600113, India
b International Centre for Theoretical Physics, P.O. Box 586, 34100 Trieste, Italy
c Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai - 400 005, India

Received 18 February 1998
Editor: L. Alvarez-Gaumé

Abstract

We consider the gravitational scattering of point particles in four dimensions, at Planckian centre of mass energy and low momentum transfer, or the eikonal approximation. The scattering amplitude can be exactly computed by modelling point particles by very generic metrics. A class of such metrics are black hole solutions obtained from dimensional reduction of $p$-brane solutions with one or more Ramond-Ramond charges in string theory. At weak string coupling, such black holes are replaced by a collection of wrapped D-branes. Thus, we investigate eikonal scattering at weak coupling by modelling the point particles by wrapped D-branes and show that the amplitudes exactly match the corresponding amplitude found at strong coupling. We extend the calculation for scattering of charged particles. © 1998 Published by Elsevier Science B.V. All rights reserved.

1. Introduction

It is expected that the quantum gravity effects would become important at energies compared to the Planck Scale. Since the gravitational coupling constant $G$ is not dimensionless, one can construct two independent dimensionless coupling constants, which for example in four space-time dimensions can be defined as $G_\parallel \equiv G_s$ and $G_\perp \equiv G_t$. Here $G_\parallel$ is the four dimensional Newton’s constant, and $s, t$ are the Mandelstam variables. Remarkably, it was shown in [1] that the full theory of quantum gravity can be split up into two independent theories with these coupling constants. Thus, a quantum gravitational regime for gravitational scattering can be envisaged when either $G_\parallel \approx 1$ or $G_\perp \approx 1$, or both. While the former signifies Planckian centre-of-mass energies, the latter implies Planckian momentum transfers. The first quantum gravitational scenario is easier to deal with because although $s$ is large, $t$ can be held fixed at a relatively small value such that $s/t \to \infty$. The impact parameter of scattering, in this case, is very large and the scattering is almost forward. This is the so called eikonal approximation where the exact two particle scattering amplitude can be computed. In practice, it is advantageous to view one of the particles as static (say $A$) and the other moving (say $B$) at almost luminal velocity past it with a large impact parameter. $A$ being static, can be suitably modelled by a metric, whose gravitational field $B$ is supposed to experience. Then one can solve the
wave equation for $B$ in this given background and obtain the scattering amplitude. Of course, the reverse process is equally valid, when the ‘shock-wave’ space-time produced by the $A$ is obtained by Lorentz boosting the metric and analysing the wave function of the slow particle in this shock-wave background. As expected, the two pictures yield the same result [2].

In the above picture, the point particles are usually modelled by Schwarzschild or Reissner-Nordström metrics, depending on whether the particles are neutral or charged. Although this seems natural in the framework of general relativity, these specific choices are certainly not mandatory. We show here that the results can be extended for a large class of generic spherically symmetric metrics. As an example, a large class of metrics arise as solutions of low energy string theory and one could model the particles by these metrics as well. The black holes which carry the NS-NS charges have already been considered in Planckian scattering which lead to interesting consequences [3]. It is shown here that these black hole metrics do not fall into the class of metrics we consider here, except in the extremal limit. Recently a new class of black hole solutions of low energy effective action of superstring theory have been found, whose weak coupling (in terms of the string coupling $g$) description consists of certain configurations of solitonic string states or D-branes, wrapped on suitable compact manifolds. Several pieces of evidence have emerged supporting this identification, the most notable being the fact that the degeneracy associated with the D-brane configuration exactly reproduces the Bekenstein-Hawking entropy [4] and open string interactions on the D-brane reproduce the Hawking radiation spectrum of these black holes [5]. Thus, while the black hole description can be used at large coupling, the D-brane description is appropriate at small coupling [4–6].

D-brane scattering has been considered in [7–12], and in particular in [7–10], it has been shown scattering amplitudes of R-R charged $p$-branes agree with appropriate D-brane scattering in ten dimensions. There are black hole solutions with a singular horizon obtained by wrapping these R-R charged $p$-branes on compact spaces. The entropy for these black holes is zero, and hence the appropriate process to check for D-brane black hole correspondence is to look at scattering amplitudes. In this paper, we show that indeed the exact eikonal scattering amplitude can be computed for wrapped D-branes at weak coupling. Moreover, this amplitude agrees with that found in the black hole picture. The agreement persists when the particles carry $U(1)$ charges also.

In the following section, we calculate the eikonal scattering phase shift in the strong coupling regime by modelling the particles by a general spherically symmetric black hole metric in four space-time dimensions. We also consider the special cases of R-R and NS-NS charged black holes. In the next section, we calculate the corresponding phase shift at weak coupling using D-$p$-branes wrapped on tori. The D-brane result is found to be independent of the brane dimensions as long as they are completely wrapped on the internal tori. Significantly, the scattering phase shift is dominated by graviton exchanges at ultrarelativistic velocities, as anticipated earlier in the calculations of [2]. Finally, we extend the calculations to include particles carrying electric charges, where too the results continue to agree.

2. Eikonal scattering at large string coupling

We will assume that the slow target particle gives rise to the following most general spherically symmetric metric in four dimensions:

$$ds^2 = -\lambda^2 dt^2 + \frac{1}{\lambda^2} dr^2 + R r^2 d\Omega^2,$$

where $\lambda$ and $R$ are functions of the radial coordinate only.

In general let the ultrarelativistic particle of charge $e$ be minimally coupled to the $U(1)$ gauge field $A_\mu$ produced by the static particle. Then its wave function $\Phi$ satisfies the covariant Klein-Gordon equation

$$\frac{1}{\sqrt{-g}} D_\mu(\sqrt{-g} g^{\mu\nu} D_\nu \Phi) + m^2 \Phi = 0,$$

$$D_\mu = \partial_\mu - ieA_\mu,$$

where $m$ and $E$ are the mass and energy of the test particle. In the spirit of [2] we will assume that $M, m \ll M_{pl}$, where $M_{pl}$ is the Planck mass. In the centre of mass frame of the particles, their energies are proportional to $\sqrt{s}$ and for the particles to have relativistic velocities at this energy $M, m \ll \sqrt{s}$. Since
in our problem $\sqrt{s} \sim M_p$, we get the above condition on the mass. Thus we will ignore all terms quadratic in $M,m$ in subsequent discussions. We substitute a solution of the form
\[
\Phi(r,t) = \frac{\Phi'}{r} e^{iE_0 \gamma_0(\theta,\phi)},
\]
and linearize the metric components at large distances as
\[
\lambda^2 = 1 - \frac{2GM}{r}, \quad R = 1 + \mathcal{O}(1/r).
\]
Here $M$ is the ADM mass of the black hole metric. Retaining terms up to order $1/r^2$, the radial equation of the $l^{th}$ partial wave is for large $l$
\[
d^2\Phi' \left[ \frac{\lambda(l+1)}{r^2} - E^2 - \frac{4G_4ME^2}{r} + 2\epsilon\mathcal{A}_0 \right] \Phi' = 0.
\]
(3)
The gauge potential is assumed to have $A_0$ as the only non-zero component. With its explicit spherically symmetric form
\[
A_0 = K/r
\]
(4)
and the identity $s = 2ME$, the above equation reduces to
\[
d^2\Phi' \left[ \frac{\lambda(l+1)}{r^2} - E^2 - \frac{2(G_4s - eK)}{r} \right] \Phi' = 0.
\]
(5)
It is straightforward to obtain the phase shift from here and the answer is [13]
\[
\delta_l = \arg \Gamma \left( l + 1 - iG_4s \right).
\]
(6)
Expanding the rhs for $l \gg 1$, we get [14]
\[
\delta_l = -(G_4s - eK) \ln l.
\]
(7)
The above phase shift resembles Rutherford scattering with the fine structure constant $\alpha$ being replaced by the effective coupling constant $-(G_4s - eK)$, which is attractive for large $s$. The phase shift can be substituted in
\[
f(s,t) = \frac{1}{2ivl}s \sum_{j=0}^{\infty} (2l+1) [e^{2i\delta_j} - 1] P_l(\cos \theta)
\]
(8)
to obtain the scattering amplitude. Using the asymptotic formula for large $l$
\[
P_l(\cos \theta) \rightarrow J_0((2l + 1) \sin \theta/2),
\]
and converting the sum into an integral as in [1], we get
\[
f(s,t) = -iv\sqrt{s} \int_0^\infty dy y^{1-2/2G_4s} J_0 \left( 2y\sqrt{s} \sin \frac{\theta}{2} \right) ,
\]
(9)
where $y = l/\sqrt{s}$ and $t = 4E^2\sin^2\theta/2$. Finally, one gets
\[
f(s,t) = \frac{-i(G_4s - eK)}{\pi t} \frac{\Gamma (1 - i(G_4s - eK))}{\Gamma (1 + i(G_4s - eK))}
\]
\[
\times \left( \frac{4}{s} \right)^{-i(G_4s - eK)}.
\]
(10)
The cross section follows:
\[
\sigma(s,t) = \frac{4(G_4s - eK)}{t^2}.
\]
(11)
The amplitude (10) exhibits the infinite set of 't Hooft poles at the values $G_4s - eK = -iN$, where $N = 1,2,\ldots \infty$. Note that for ordinary particles, $eK \ll 1$ and the electromagnetic contribution to the scattering is suppressed at Planckian energies. In other words, gravity assumes the role of the dominant interaction at the Planck scale.

The linearizations of the functions $\lambda$ and $R$ are valid for spherically symmetric metrics of general relativity as well as black holes carrying Ramond-Ramond charges. Examples in four dimensions are [15]
\[
ds^2 = -f^{-1/2}hdt^2 + f^{1/2} \left( h^{-1}dr^2 + r^2d\Omega^2 \right),
\]
(12)
where $h = 1 - r_o/r$, $f = \prod_{i=1}^{4} (1 + r_i/r)$ and the horizon is at $r_o$. This black hole metric arises when three distinct five branes of $\text{M-theory}$ intersect along a line which is then wrapped on a circle. The parameters $r_i, i = 1,\ldots,4$ are related to the four $U(1)$ charges carried by the black hole, three of which are proportional to the number of the three different 5-branes, while the fourth is proportional to the Kaluza-Klein momentum along the intersection line. The ADM mass of this black hole is $M = (\sum_{j=1}^{4} r_j + 2r_o)/4G_4$. For solutions obtained by wrapping BPS saturating
fundamental strings or R-R charged p-branes, the
the asymptotic value of the dilaton field. There is
mass of the p-brane. The horizon is at \( r = 0 \), and
thus is singular in nature. But this metric is well
behaved at large distances and gives the exact scat-
tering amplitude as shown in the above calculation.

On the other hand, certain black holes carrying
NS-NS charges have non-extremal metric of the
form [17]
\[
ds^2 = -\left(1 - \frac{2GM}{r}\right)dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\frac{dr^2}{1 - \frac{\alpha}{Mr}} + r^2\left(1 - \frac{\alpha}{Mr}\right)d\Omega^2,
\]
where \( \alpha = Q^2e^{-2\phi_0} \), \( Q \) being the electric charge and
\( \phi_0 \), the asymptotic value of the dilaton field. There is a
curvature singularity at the horizon \( r = \alpha/M \),
which expands without limit for vanishing masses. It
Can be seen that the corresponding \( R(r) \) (Eq. (1))
here cannot be linearised in the asymptotic \( r \to \infty \)
region, except in the extremal limit. Thus, the eikonal
scattering can be computed with these metrics only
in the extremal limit [3].

3. Eikonal scattering at small string coupling

In this section, we will compute the eikonal phase
shift at weak string coupling, when the D-brane picture
is appropriate. We develop a general formul-
ism for the scattering of wrapped D-branes before
specialising to four dimensions.

Consider a D-p-brane moving with a relative ve-
clocity \( v \) with respect to a D-l-brane in 10 space-time
dimensions. They are separated by a large transverse
distance \( b \). We assume \( l \leq p \) and that none of the
coordinate directions of the D-l-brane are orthogonal
to those of the D-p-brane. Apart from the direction of
velocity and the time coordinate, the end points of an
open string ending on the two branes satisfy
either Neumann (N) or Dirichlet (D) boundary condi-
tions. We denote as \( NN \) the number of string coordi-
nates which satisfy N condition at both the ends.

Similarly \( ND \) and \( DD \). Evidently, \( DD = 8 - p \), \( NN = l \)
and \( ND = p - l \). The scattering phase shift be-
tween these two branes are given by the one loop
vacuum superstring amplitude [7,8]
\[
\delta(b) = \frac{1}{2} \int_0^\infty \frac{dNNk}{d(2\pi)^{1+\alpha'}} \sum_i \int_0^\infty \frac{dt}{t} e^{-2\pi'\alpha'\alpha}\left(1 + M_i^2\right),
\]

where
\[
M_i^2 = \frac{b^2}{4\pi'\alpha} + \frac{1}{\alpha} \sum (\text{oscillators}).
\]
The oscillator sum and the integral finally yields
\[
\delta(b) = \frac{1}{4\pi} \int_0^\infty \frac{dt}{t} \left(8\pi'\alpha'\right)^\alpha e^{-4\pi'\alpha'/2}\left(1 - B \times J\right),
\]

where \( B, J \) are the bosonic and fermionic contri-
butions to the oscillator sum in the in the one loop open
superstring amplitude given by [8]
\[
B = f_1^{(NN+DD)}(q)f_4^{NN}(q)\frac{\Theta^1(0\mid it)}{\Theta^1(\epsilon\mid it)},
\]
\[
J = \frac{1}{2} \left[f_2^{NN+DD}(q)f_3^{ND}(q)\frac{\Theta^1(\epsilon\mid it)}{\Theta^1(0\mid it)} + f_3^{NN+DD}(q)f_2^{ND}(q)\frac{\Theta^1(\epsilon\mid it)}{\Theta^1(0\mid it)} \right],
\]

where \( q \equiv e^{-\pi t} \),
\[
f_1(q) = q^{1/12} \prod_{n=0}^{\infty} (1 - q^n),
\]
\[
f_2(q) = \sqrt{2} q^{1/12} \prod_{n=0}^{\infty} (1 + q^n),
\]
\[
f_3(q) = q^{-1/24} \prod_{n=0}^{\infty} (1 + q^{2n-1}),
\]
\[
f_4(q) = q^{-1/24} \prod_{n=0}^{\infty} (1 - q^{2n-1}).
\]
The rapidity \( \epsilon \) is defined as \( \tanh \pi \epsilon = v \). Note that the
last term in \( J \) comes from summation of the
\( NS(-1)^F \) sector and contributes only when \( ND = 0 \).
This term is due to R-R exchange, and we will concentrate on this term in the next section when we look at charged particle amplitudes.

Now, to compare the D-brane results with the results on the black hole side, we have to compactify the branes on suitable compact manifolds, such that in the non-compact space-time they look like point particles. For simplicity, we compactify on a c dimensional torus (with \( p \leq c \)), such that the resultant noncompact space time is a \((10-c)\) dimensional with a Lorentzian signature. We take the range of each coordinate of the torus to be \( L_i, i = 1, \ldots, c \). Thus, the volume of the torus is \( V = \prod L_i \). To obtain the phase shift for these wrapped D-branes, the formula (16) has to be modified. For each compactified NN direction, the momentum integral is replaced by a discrete sum in the one loop amplitude. In other words the momentum integral is restricted to the remaining non-compact NN directions, and and a factor of \( \Theta_4(0, \pi \alpha'/t/L^2) \) is inserted in the integrand for each compactified direction. Similarly, for each compact DD direction, a sum over winding modes is introduced, resulting in a factor of \( \Theta_4(0, i\pi l^2/2 \pi \alpha') \) in the integrand [9]. Since all the NN=1 coordinates are compactified and there are \( c-p \) compact DD coordinates, the final result is

\[
\delta(b) = \frac{1}{4\pi} \int_0^{\infty} \frac{dt}{t} e^{-b^2 t/2 \pi \alpha'} \times \left[ \prod_{i=1}^{l} \Theta_4(0, 8 \pi \alpha'/t/L^2) \right] \times \left[ \prod_{j=p+1}^{c} \Theta_4(0, i\pi l^2/2 \pi \alpha') \right] \cdot (B \times J).
\]

\( (23) \)

Now, large impact parameter \((b \to \infty)\), scattering is dominated by the exchange of massless closed string states, for which it is sufficient to restrict the integrand in the regime \( t \to 0 \) [10,18]. Using the relevant formulae given in [9,19], we get

\[
\Theta_4(0, 8 \pi \alpha'/t/L^2) \to \frac{L}{2^{3/2} \pi^2 \alpha'} \frac{1}{\sqrt{t}}.
\]

\( (24) \)

\[
\Theta_4(0, i\pi l^2/2 \pi \alpha') \to \frac{2^{1/2} \alpha}{L} \frac{1}{\sqrt{t}}.
\]

\( (25) \)

\[B \to 2^{-p-1/2} \frac{1}{2^{3/2} \pi^2 \alpha'} \frac{e^{-\pi \epsilon^2}}{\sinh \pi \epsilon}, \quad (26)\]

\[J \to 4 e^{-2\pi/3l} e^{\pi \epsilon^2}, \quad (27)\]

and the phase shift becomes

\[
\delta(b) = \Lambda \kappa(\epsilon) \int_0^{\infty} \frac{dt}{t} e^{-b^2 t/2 \pi \alpha'} \frac{t^3}{t^{1/2}}, \quad (28)\]

\[\Lambda = \frac{2^{3/2} \pi^2 e^{-(p+1) \xi - (p+1) \eta} \prod_{i=1}^{l} L_i}{2^{p+l} \prod_{i=1}^{l} L_i} \quad (29)\]

\[\kappa(\epsilon) = \frac{2 - (p - l)/2 + \sinh^2 \pi \epsilon - 2 \delta_{p,l} \cosh \pi \epsilon}{\sinh \pi \epsilon} \quad (30)\]

Note that the integral in the above expression is independent of \( p \) and \( l \) and depends only on \( c \). That is, it is the same for branes of arbitrary dimensions for a given compactification. Thus, scattering phase shifts for all \( p, l \) can be calculated from the above expression provided the branes completely wrap on the internal torus. Also, as expected, \( K(\epsilon = 0) = 0 \) for \( p = l \) or \( p-l = 4 \). This is the familiar no-force condition for BPS states.

However, for our present purposes, we specialise to the case of \( c = 6 \), i.e. scattering in 4 dimensional non-compact space time. Then the integral over \( t \) in (28) simply yields a factor \(-2 \ln(b/\sqrt{2 \pi \alpha'})\). In addition, we make the ultrarelativistic approximation \( v \to 1 \), \( \epsilon \to \infty \), such that

\[\kappa(\epsilon) \to \frac{1}{2} e^{\pi \epsilon}.\]

Then Eq. (28) becomes

\[
\delta(b) = \frac{1}{2} \ln \frac{b}{\sqrt{2 \pi \alpha'}} \quad (31)\]
Rewriting this amplitude in terms of the masses of the branes given by
\[ m_{p(l)} = \frac{\rho(l)}{g(2\pi)^{d(l)}\sqrt{\alpha'}} \prod_{i=1}^{p(l)} L_i \]  
and the Newton’s constant in \((10 - c)\) dimensions (for \(c = 6\))
\[ G_{10-c} = \frac{8\pi^6 g^{-\frac{3}{4}}}{V} \]  
yields
\[ \delta(b) = -G_4 m_p m_1 e^{\pi\epsilon} \ln \frac{b}{\sqrt{2\pi\alpha'}}. \]  
Finally, using \(s = m_1 m_p \exp(\pi\epsilon)\), which simply expresses the relativistic transformation of energy, and \(l = bE\), the amplitude takes the form
\[ \delta(b) = -G_4 s \ln \frac{l}{E\sqrt{2\pi\alpha'}}. \]  
(the factor \(E\sqrt{2\pi\alpha'}\) is irrelevant as it does not appear in the scattering amplitude and the cross section, and will simply be dropped).

Thus not only is the phase shift exactly calculable in the weak coupling regime, comparison with (7) (for \(e = 0\)) shows that it perfectly agrees with that calculated in the strong coupling regime, implying that there is no discontinuity in the point particle scattering amplitude as one tunes the string coupling, for arbitrary masses of particles. Moreover, as expected, the gravitational interaction dominates overwhelmingly over gauge interactions.

4. Inclusion of charge

To extend the results of the previous section to include gauge interactions, the charge interaction term proportional to \(\cosh\pi\epsilon\) has to be retained in the kinematical factor (30). This vanishes unless \(p = l\), since branes of different dimensions do not couple to each other via gauge fields. Thus for \(p = l\)
\[ \kappa(e) \rightarrow \frac{1}{2}(e^{\pi\epsilon} - 4). \]

As before, introducing \(m_p\) and \(G_4\) gives
\[ \delta(b) = -G_4 m_p^2 (e^{\pi\epsilon} - 4) \ln \tilde{l} \]  
\[ = -(G_4 s - 4G_4 m_p^2) \ln \tilde{l}. \]

Let us now consider the phase shift obtained using the black hole background, Eq. (7). Here, inclusion of charge shows that the coupling constant is given by \(G_5 - eK\). To determine \(K\), we look at the \(p + 1\) form potentials due to the static brane to which the relativistic \(p\) brane couples. In \(10\)-dimensions the asymptotic value of the \(p + 1\) potential is given by [10]
\[ A_{p+1} = \frac{q_p}{r^{7-p}} dt \wedge dx \wedge \ldots \wedge dx^n, \]  
where
\[ q_p = 2^{5-p} \frac{g_4}{4} \sqrt{\frac{5-p}{2}} \pi^{\frac{5-p}{2}} \Gamma\left(\frac{7-p}{2}\right). \]

The effective \(U(1)\) potential due to the brane living in \(R^{10-c} \times T^c\) is obtained in two steps: Firstly, a Kaluza Klein reduction is performed on the \(10\)-dimensional \(p + 1\) potential to obtain a one form potential in \(10 - p\) dimensions due to the brane completely wrapped on \(T^p\) [20]:
\[ A_1 = \frac{q_p}{r^{7-p}} dt. \]  
Secondly, for the remaining \(c - p\) compact directions, which are transverse to the brane vertical reduction is performed. For this, we stack the Kaluza-Klein reduced configurations in the compact directions transverse to the brane and go to the continuum limit by integrating over the latter, resulting in one form potentials in \(10 - c\) dimensions [20]
\[ A_1 = \frac{q_p}{V} \prod_{i=1}^{p} L_i \int_{\alpha}^{\infty} \frac{d^{c-p} r_x}{\left[r^2 + r_x^2\right]^{2}} \]  
\[ = \frac{q_p}{V} \prod_{i=1}^{p} L_i \int d\Omega_{c-p-1} \int_{0}^{\infty} \frac{r_x^{c-p-1} dr_x}{\left[r^2 + r_x^2\right]^{2}} \]  
\[ = \frac{q_p}{V} \prod_{i=1}^{p} L_i \int d\Omega_{c-p-1} \int_{0}^{\infty} \frac{r_x^{c-p-1} dr_x}{\left[r^2 + r_x^2\right]^{2}}. \]
where \( r_\perp \) refers to the transverse distance from the 0-brane, and \( \Omega_{e-p-1} \) is the volume of the unit \((e-p-1)\) sphere, given by

\[
\Omega_{e-p-1} = \frac{2\pi^{e-p}}{\Gamma\left(\frac{e-p}{2}\right)}.
\]

The integral is elementary, and when expressed in terms of the \( p \)-brane mass and the Newton’s constant given Eqs. (32) and (33) respectively, we get

\[
A_1 = \frac{4G_{10-e} \Gamma\left(\frac{7 - e}{2}\right)}{\pi^{\frac{7 - e}{2}} m_p dt}.
\]  

(43)

Comparing with the expression (4), we conclude that

\[
K = \frac{4G_{10-e} \Gamma\left(\frac{7 - e}{2}\right)}{\pi^{\frac{7 - e}{2}} m_p}.
\]

Now, the charge of the moving brane is \( e = m_p \), by the BPS condition. Consequently, for \( c = 6 \), the scattering phase shift \( \delta(\theta) \) will be modified as

\[
\delta(\theta) = -(G_4 \delta - 4G_4 m_p^4) \ln \frac{1}{\theta}.
\]  

(44)

which precisely agrees with the phase shift (37) obtained in the D-brane side.

5. Discussions

We have shown that the eikonal scattering amplitude obtained by modelling the point particles as black holes is exactly reproduced by eikonal scattering of wrapped D-branes. In this regime only the gravitational field at infinity is probed (as the black hole metric is linearised) and the details of the metric are not realised. So, one should study the corrections to the eikonal phase shift, as the impact parameter and velocity are tuned to smaller values, and see whether the details begin to emerge [12,21]. Our results are independent of the dimensions of the brane but the kinematical factor \( \kappa(e) \), and hence \( \delta(b) \) ceases to be independent of \( p \) and \( l \) once one relaxes the condition \( e \to \infty \).

We know that the type II B string theory has S-duality group as SL(2,Z) which relates fundamental strings to the D-strings. Under this S-duality operation, gravitons are left invariant and NS-NS charged fields become R-R charged fields. Our D-brane calculation for gravitational exchange (dominant term) matches the leading order term (eikonal limit) in the scattering amplitude for fundamental strings [22]. This confirms the S-duality symmetry. Moreover, we have obtained the subdominant term (eikonal limit) due to R-R charged field exchange between the two D-p-branes. Invoking the S-duality, we can say that this must be also be the amplitude for NS-NS gauge field exchange in fundamental string scattering. In [22], corrections to the eikonal fundamental string-string graviton exchange amplitude were calculated and shown to be order \( 1/\tau^2 \). It will be interesting to see whether the D-brane scattering amplitude gives the same.

Though the fundamental string and the D-brane scattering amplitudes are same, the relevant energy scales for eikonal scattering are different for the two. For the D-p-branes to be relativistic, the condition

\[
E \gg m_p
\]  

(45)

must be imposed on their energy. Using the expression for the brane mass, we get

\[
E \gg \frac{L_p}{g \sqrt{\alpha'}^{p+1}}.
\]

As is evident, in 10-dimensions, the mass of the D-0 brane is much larger than the Planck mass \( M_{pl} \) and the string mass \( m_s \sim 1/\sqrt{\alpha'} \). Hence for condition (45) to be realised, energies relevant for D-0 brane eikonal scattering has to be much larger than both \( M_{pl} \) and \( m_s \). In other words for D-0 branes to become relativistic, we need to consider regimes where \( g \sqrt{\alpha'} \gg 1 \) (In the c.m. frame \( s = E^2 \)). (Note that for non-perturbative effects of M-theory to become important, we need to have \( g \sqrt{\alpha'} \to 1 \) [18], and we are not probing that regime.) On the other hand, for fundamental strings, \( s \alpha' \gg 1 \) gives the eikonal limit, as considered in [22]. Following [23] we take \( 1/g \sqrt{\alpha'} \) as the energy scale in our problem. and the condition on the compactification volume, from Eq. (44) is

\[
V_p \ll \sqrt{\alpha'}^p.
\]
which implies that the compactification radius should be sufficiently small compared to the string scale. Note that if we had used $1/\sqrt{\alpha'}$, as in [22] the conditions on compactification lengths would have become $g$ dependent, and difficult to interpret.

Another interesting observation is that though the D-brane scattering amplitude includes all long range closed string exchanges, only the graviton exchange dominates in the above kinematical regime. This was anticipated in the black hole calculation by 't Hooft and the weak coupling calculation vindicates this. Our weak coupling calculations can be generalised to higher dimensions. For example, in five non-compact dimensions, it is easy to see that the phase shift goes as $1/r$, which is the Green’s function for three transverse dimensions.

One can try to examine more sophisticated compactifications e.g. on $K3$ to see whether similar conclusions hold for those situations also [25]. We hope to report on it in the near future.

Acknowledgements

S.D. would like to thank C. Bachas, F. Hussain, R. Iengo and K. S. Narain for useful discussions and the International Centre for Theoretical Physics, Trieste for hospitality where part of this work was done. A.D. would like to thank Theoretical Physics Group, Tata Institute of Fundamental Research, Mumbai for hospitality during the course of this work and P. Majumdar for useful discussions.

References


This was also observed in [24]
Strings and D-branes with boundaries

C.S. Chu a, P.S. Howe b, E. Sezgin c, 1

a International School for Advanced Studies (SISSA), Via Beirut 2, 34014 Trieste, Italy
b Department of Mathematics, King’s College, London, UK
c Center for Theoretical Physics, Texas A&M University, College Station, TX 77843, USA

Received 7 February 1998
Editor: M. Dine

Abstract

The covariant field equations of ten-dimensional super D-branes are obtained by considering fundamental strings whose ends lie in the superworldsurface of the D-brane. By considering in a similar fashion D p-branes ending on D(p + 2)-branes we derive equations describing D-branes with dual potentials, as well as the vector potentials. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

In a recent paper [1], an open supermembrane ending on an M-fivebrane was studied. The worldvolume of the M-fivebrane was taken to be a super-submanifold, M, of the eleven-dimensional target superspace, M. The supermembrane action is an integral over a bosonic three dimensional worldvolume, Σ, with its boundary ∂Σ embedded in the supermanifold M, such that

\[ \partial \Sigma \subset M \subset \Sigma \]  

(1)

It was shown that the \( \kappa \)-symmetry of the total action implies (1) the eleven-dimensional supergravity equations, (2) a constraint on the embedding of \( M \) in \( \Sigma \) and (3) a constraint on a modified super 3-form field strength, \( H \), on the superfivebrane worldvolume. The superembedding constraint and the \( H \)-constraint completely determine the superfivebrane equations of motion, although the \( H \)-constraint can be derived from the embedding constraint in this case.

In this paper, this result is extended to the following configurations of branes:
1. Fundamental type II strings ending on D-branes
2. Type II Dp-branes ending on D(p + 2)-branes

The target space is (10|32) dimensional type IIA or type IIB superspaces. We use the notation \( (D|D') \), where \( D \) is the real bosonic dimension and \( D' \) is the real fermionic dimension of a supermanifold. The embedded supermanifold \( M \) has dimension \( 2|16 \) in case (1) and \( (p + 1|16) \) for case (2).

In all cases we find that the \( \kappa \)-symmetry of the total system implies the ten-dimensional type IIA or type IIB supergravity equations and a constraint on the embedding of \( M \) in \( \Sigma \). In addition, in case (1) we find a constraint on a modified super 2-form field strength on \( M \) defined as

\[ F = dA - B \, , \]  

(2)

1 Research supported in part by NSF Grant PHY-9722090.
where $A$ is the super 1-form potential on $M$, and $B$ is the pullback of the target space NS-NS super 2-form $M$. We will use the same letter to denote the target space and worldsurface superforms, since it should be clear from the context whether a pullback is required. The superembedding constraint and the $\mathcal{F}$-constraint determine completely the dynamics of the D-brane on which the fundamental string ends. At the linearised level, these constraints are shown to be precisely the dimensional reduction of the ten-dimensional Maxwell superspace constraints.

In case (2) the construction leads naturally to the introduction of a $p$-form potential on $M$ in addition to the usual one-form potential. The field strength forms corresponding to these potentials are essentially dual to one another, so that the dual versions of D-branes are automatically generated by this method. Again, constraints on the field strengths are derived and imply the equations of motion for the D-brane when the embedding condition is taken into account.

In the case of M-fivebrane, the 3-form $H$ was introduced in [2,3] for convenience in describing the field equations and it was shown that the $H$-constraint is a consequence of the superembedding condition [3,4]. In [2], it was observed that the analogue of the $H$-constraint arises in the description of various superbranes and, in particular, the super D-branes naturally accomodate an $\mathcal{F}$-constraint, where $\mathcal{F}$ is a modified two-form field strength. It was also noted in [2] that for certain superbranes, e.g. the D6-brane, the $\mathcal{F}$-constraint is needed to put the theory on-shell.

In the approach presented in [1], both the superembedding condition and the $H$-constraint arise naturally from the requirement of $\kappa$-symmetry. Similarly, here we will show that both the superembedding condition and the $\mathcal{F}$-constraint arise naturally from the considerations of $\kappa$-symmetry of suitable open branes ending on D-branes.

2. Fundamental type II strings ending on D-branes

In this section, we consider the fundamental type II strings ending on D-branes. The string worldsheet is bosonic. We will take its boundary, however, to lie in a bosonic submanifold of a supermanifold $M$ of dimension $(p + 1)[16]$, which in turn is a submanifold of a target space $M$ of dimension $(10)[32]$. We use the notations and conventions of [2]. In particular, we denote by $z^M = (x^m, \theta^\alpha)$ the local coordinates on $M$, and by $A = (a_\alpha, \bar{\alpha})$ the target tangent space indices. We use the ununderlined version of these indices to label the corresponding quantities on the worldsurface. The embedded submanifold $M$, with local coordinates $z^M = (x^m, \theta^\nu)$, is given as $z^M(\mathbf{z})$.

We shall consider type IIA and type IIB superspaces. The fermionic coordinates consist of two Majorana-Weyl spinors. In type IIA superspace, these spinors carry opposite chiralities which can be combined into a single 32 component Majorana spinor, while in type IIB superspace they are of the same chirality. We will use the fermionic index $a$ in both cases, though in type IIB superspace it is understood to be a composite index of Majorana-Weyl spinor index and an $SO(2)$ doublet index, acted on by a direct product of chirally projected $\Gamma$-matrices and $SO(2)$ matrices. Further details of our notation and conventions are given in the Appendix.

2.1. Constraints from $\kappa$-symmetry of the open string

We consider the following action for the total system of a type II open string ending on a D-brane (with the target metric taken to be in the the Einstein frame),

$$S = -\int_{\Sigma} d^2 \xi \left( \sqrt{-g} + \epsilon^{ij} B_{ij} \right) + \int_{\partial \Sigma} d \tau A, \quad (3)$$

where $\xi^i (i = 0,1)$ are the coordinates of the string worldsheet $\Sigma$, $\tau$ is the coordinate on the boundary $\partial \Sigma$. We will take both ends of the string to lie on a $Dp$-brane supermanifold $M$ of dimension $(p + 1)[16]$. $A$ is the pullback to $\partial \Sigma$ of a super one-form
defined on $M$. The induced metric $g_{ij}$, and the pullbacks $B_{ij}, A$ are defined as:

\[ B_{ij} = E_i \varepsilon E_j \Delta A_{AB}, \]

\[ A = E_i \varepsilon A_{ij}, \]

\[ g_{ij} = E_i \varepsilon E_j \Delta \eta_{ab}, \tag{4} \]

where $\eta_{ab}$ is the Minkowski metric in ten dimensions, and

\[ E_i \Delta = \delta_i z^A E_M \Delta A, \]

\[ E_r \Delta = \delta_i z^A E_M \Delta A, \tag{5} \]

where $E_M \Delta$ is the target space supervielbein and $E_M$ is the worldsurface supervielbein. We note the useful relation

\[ (\Gamma_{ij})_{\Delta} = \frac{1}{2\sqrt{-g}} \varepsilon^{ij}(\gamma_{ij} P)_{\Delta}, \tag{9} \]

where

\[ P = \begin{pmatrix} \Gamma_{11} & (IIA) \\ \sigma_3 & (IIB) \end{pmatrix} \tag{10} \]

Note that the $\Gamma_{11}$ acts on a 32 component Majorana index, while $\sigma_3$ acts on the $SO(2)$ doublet index of the two 16-component Majorana-Weyl spinors of same chirality.

The boundary $\kappa$-transformations will be taken to be of the form as in [1], namely

\[ \delta_\kappa z^a = 0, \]

\[ \delta_\kappa z^a = \frac{1}{2} \kappa \xi (1 + \Gamma_{(p+1)}) z^a \text{ on } d\Sigma, \tag{11} \]

where the matrix $\Gamma_{(p+1)}$ is defined by

\[ E_a \Delta z^a = \frac{1}{2} (1 + \Gamma_{(p+1)}) \Delta \tag{12} \]

The matrix $E_a \Delta z^a$ is obtained from $E_A \Delta$ which is the inverse of $E_A \Delta$. For more details, see [3,4].

The vanishing of the terms on $\Sigma$ imposes constraints on the torsion super two-form $T$, and the super three-form $H = dB$, such that they are consistent with the equations of motion of the ten-dimensional type II supergravities [6].

The constraints which follow from $\kappa$-symmetry on $\Sigma$ are

\[ T_{\delta \beta} = -i(\Gamma \Delta g^a \Delta), \tag{13} \]

\[ T_{\delta \beta} = \delta_i z^A E_M \Delta A, \]

\[ H_{\delta \beta} = i(\Gamma \Delta Q)_{\delta \beta}, \tag{14} \]

where

\[ Q = \begin{pmatrix} \Gamma_{11} & (IIA) \\ \sigma_3 & (IIB) \end{pmatrix} \tag{15} \]

and $\chi_a$ is a spinor superfield proportional to the dilaton superfield of the supergravity background.

The remaining variations are on the boundary. Proceeding exactly as in [1], we learn that they vanish provided that the following two constraints are satisfied:

\[ E_a \Delta = 0, \tag{16} \]

\[ \mathcal{F}_{AB} = 0. \tag{17} \]

Here,

\[ dA = B, \tag{18} \]

is the modified 2-form superfield strength which satisfies the Bianchi identity

\[ d\mathcal{F} = -H. \tag{19} \]

There will also be a mixture of Dirichlet and Neumann boundary conditions from the requirement that the action be stationary when the string field equations hold. These can be derived straightforwardly as
in the case of the open supermembrane which has been discussed in detail in [1].

2.2. Solution of the linearised constraints

In this section, we shall analyse the embedding condition (16) and the the $\mathcal{F}$-constraints (17) in order to extract the equation of motion for the D-brane worldvolume fields. To determine the field content, it is sufficient to study the linearised constraints in flat target space limit.

The supervielbein for the flat target superspace is,

\[ E = dx - i \frac{1}{2} d\theta \Gamma^a_{\alpha \beta} \theta^\beta. \]

Let us choose the physical gauge,

\[ x^a = x^a(x, \theta), \]

\[ \theta^\alpha = \theta^\alpha(x, \theta) \]

and take the embedding to be infinitesimal so that $E_i^{\alpha} \partial_M$ can be replaced by $D_a = (\partial_a, D_a)$ where $D_a$ is the flat superspace covariant derivative on the worldsurface, provided that the embedding constraint holds. In this limit the embedding matrix is:

\[ E_a^b \rightarrow \begin{cases} \delta_a^b \\ \partial_a X^b \end{cases}, \]

\[ E_a^b \rightarrow \begin{cases} 0 \\ D_a X^d - i(\Gamma^d)_{a\beta} \theta^\beta \end{cases}, \]

\[ E_a^\beta \rightarrow \begin{cases} 0 \\ \partial_a \theta^\beta \end{cases}, \]

\[ E_a^\beta \rightarrow \begin{cases} \delta_a^\beta \\ D_a \theta^\beta \end{cases}, \]

where

\[ X^d := x^d + \frac{i}{2} \theta^a(\Gamma^d)_{a\beta} \theta^\beta. \]

Using this in the embedding condition (16) we find, at the linearised level,

\[ D_a X^d = i(\Gamma^d)_{a\beta} \theta^\beta. \]

The Bianchi identity (19) in component form is,

\[ D_A \mathcal{F}_{BC} + T_{[AB} \mathcal{F}_{|EC]} = -\frac{i}{2} E_c \mathcal{F}_{AB} \mathcal{H}_{BC}. \]

Linearising this equation using (22), we find that the $ABC = (\alpha \beta \gamma)$ component of this identity is satisfied automatically, while the $\beta(\gamma \alpha)$, $\alpha(\beta \gamma)$ and $\alpha(\beta \beta)$ components give

\[ D_\gamma \mathcal{F}_{ab} = -2i \partial_\gamma \theta^\alpha (\Gamma_{\beta\gamma})_{\gamma a}, \]

\[ \partial_\gamma \mathcal{F}_{ab} = 0, \]

\[ (\gamma^a)_{\alpha \beta} \mathcal{F}_{bc} = i(\Gamma_c)_{\alpha \beta} + (\Gamma_c)_{\alpha \beta} \partial_\gamma X^d + 2 D_\alpha \theta^\beta (\Gamma_c Q)_{\beta \gamma}. \]

Our strategy is to interpret these equations, together with (24), as the dimensional reduction of the $(N = 1)$ ten-dimensional super Maxwell system to $(p + 1)$ dimensions. The relation between $\mathcal{F}$ and the non-covariant $F = dA$ follows from (18). In component form, (18) reads

\[ \mathcal{F}_{AB} = F_{AB} - E_B \mathcal{F}_{A} \Delta B_{AB}. \]

These relations are:

\[ \mathcal{F}_{ab} = F_{ab}, \]

\[ \mathcal{F}_{a\beta} = F_{a\beta}, \]

\[ \mathcal{F}_{aa} = F_{aa} - i(\Gamma a)_{a\alpha} \theta^\alpha. \]

Using (17), the constraints (26), (27) and (30)-(32) can be summarized as

\[ F_{ab} = 0, \]

\[ F_{aa} = i(\Gamma a)_{a\alpha} \theta^\alpha, \]

\[ D_\gamma F_{ab} = -2i \partial_\gamma \theta^\alpha (\Gamma_{\beta\gamma})_{\gamma a}, \]

\[ \partial_\gamma (F_{ab}) = 0. \]

It is now easy to combine these and (24) to obtain ten dimensional master constraints. To do this, we first define a ten dimensional vector superfield $A^\alpha$, and a spinor superfield $\lambda^\alpha$. \(^3\)

\[ A^\alpha, iX^d \rightarrow A^\alpha, \]

\[ \theta^\alpha \rightarrow \lambda^\alpha, \]

where the index $\alpha$ now labels a sixteen component

\(^3A\) is imaginary due to our choice (13) of the torsion.
Majorana-Weyl spinor in ten dimensions. With these definitions, the constraints (24) and (34) combine to
\[ F_{\alpha} = (\sigma_2)_{\alpha\beta} \lambda^\beta, \]
where the \( \sigma \) matrices are the ten dimensional chiral matrices. This constraint, together with (28), (33), (35), and (36) are precisely the superspace constraints of ten dimensional super Maxwell system that satisfy the ten dimensional Bianchi identity \( dF = 0 \). In particular, the constraint (28) is the dimensional reduction of the ten dimensional Bianchi identity
\[ D_{(a} F_{b)} = \frac{i}{2} (\sigma_2)_{\alpha\beta} F_{\alpha\beta} = 0, \]
for \( c = c \). The other component of this equation, i.e. for \( \bar{c} = c' \), is also satisfied, thanks to the supersymmetry algebra,
\[ \{ D_a, D_b \} = i (\sigma_2)_{\alpha\beta} \delta_{a\beta}^c. \]
In doing these calculations, we have used the properties of the \( T \)-matrices, given in the Appendix, to set the first term on the right hand side of (28) equal to zero.

We conclude that the linearised versions of our two master constraints (17) and (16) describe precisely the dimensional reduction of the ten-dimensional super Maxwell system to \( D_p \)-brane worldvolume. We expect [2,5] that the full constraints (16) and (17) imply the full field equations that follow from the super D-brane actions of [7–11].

3. \( D_p \)-brane ending on a \( D(p+2) \)-brane

In this section we study an open \( D_p \)-brane ending on a \( D(p+2) \)-brane and show that it naturally gives rise to a dual potential on the worldsurface of the \( (p+2) \)-brane \( M \).

3.1. The action and \( \kappa \)-symmetry constraints

The action for a type II\( \kappa \) symmetric D-brane has been constructed in [7–11]. \( \kappa \) symmetry implies constraints for the target superspace torsion, the NS-NS three form field strength and the RR field strength. For an open \( D(p+2) \)-brane, we propose the action (here we take the target metric to be in the string frame)
\[ S = \int \left( - e^{-\phi} \sqrt{-\det \left( g_{ij} + \mathcal{F}_{ij} \right)} + C e^{\mathcal{F}} + m \omega_{p+1} \right) + \int \omega_p, \]
where
\[ \mathcal{F} = dA - B, \]
and \( \omega_{p+1}(A, dA) \) is the Chern-Simons form present for even \( p \) in a massive IIA background, with \( m \) being the mass parameter. We define this form by the relation
\[ d\omega_{p+1} (A, dA) = (e^{dA})_{p+2}. \]
In (42), \( B \) represents the pullback of the target space super two-form to the bosonic \( D_p \)-brane worldvolume. The potential \( A_p \) is identified with the pullback onto the bosonic boundary \( \partial \Sigma \) of a \( p \)-form potential living on the \( D(p+2) \)-brane superworldvolume. Furthermore, it is assumed that the pullback of the field strength \( \mathcal{F}_p = dA_1 \) \( - B \) defined on the superworldvolume of the \( D(p+2) \)-brane onto \( \partial \Sigma \) coincides with \( \mathcal{F} \) for the \( D_p \)-brane restricted to the (bosonic) boundary. Thus we have both a one-form potential \( A_1 \) and the dual \( p \)-form potential \( A_p \) on the \( (p+2) \)-brane. For simplicity, we will use in the following the same symbol \( \mathcal{F} \) to denote both the two-form field strengths on the \( D_p \)-brane and on the \( D(p+2) \)-brane. It should be clear from the context which one is being referred to.

For later reference, we record here the definitions of the target space RR field strengths:
\[ G = dC - CH + m e^B, \]
where
\[ H = dB. \]
We also record the Bianchi identity
\[ dG = GH, \]
Note that the \( m \)-dependent terms have cancelled. We use the superspace conventions of [12] according to which the exterior derivative acts from the right.
The field strengths $\mathcal{F}, G, H$ are invariant under the gauge transformations

$$
\delta \Lambda = \lambda, \quad \delta B = d \lambda, \quad \delta C = -m \lambda \epsilon^B.
$$

(47)

where $\lambda$ is a target space super one-form gauge parameter. The field strength $G$ is also invariant under the gauge transformation

$$
\delta \Lambda = 0, \quad \delta B = 0, \quad \delta C = e^B d \mu.
$$

(48)

where $\mu$ is a target space superform of appropriate rank.

Since the gauge variation of the Chern-Simons form has the form

$$
\delta \omega_{p+1} = \lambda \epsilon^{p+1} + dX_p^1,
$$

(49)

for some $p$-form $X_p^1(\lambda, A, dA)$ defined by this equation, the action (41) is invariant under the gauge transformations (47) and (48), provided that $A_p$ transforms as

$$
\delta A_p = -m X_p^1 - \mu \epsilon^{p+1}.
$$

(50)

Next, we turn to the discussion of $\kappa$-symmetry. One can verify that, under a $\kappa$-symmetry transformation, the vanishing of the variations on $\Sigma$ impose constraints on the superscription $T$, the NS-NS field strength $H$ and the RR field strengths $G$ [7-11] such that they are consistent with the field equations of the type II supergravities.

The remaining variations are on the boundary, and they take the form

$$
\delta S = \int_{\Sigma} i_{\kappa} \mathcal{F}_{p+1},
$$

(51)

where the modified field strengths $\mathcal{F}_{p+1}$ for the worldvolume potentials $A_p$ are

$$
\mathcal{F}_{p+1} = dA_p + (Ce^\varphi)_{p+1} + m \omega_{p+1}.
$$

(52)

They satisfy the Bianchi identity

$$
d\mathcal{F}_{p+1} = (G e^\varphi)_{p+2}.
$$

(53)

Observe that the $m$-dependent terms have cancelled. Using $i_{\kappa} \mathcal{F} = 0$, the vanishing of (51) implies that

$$
\mathcal{F}_{\alpha \beta \cdots} = 0.
$$

(54)

In addition, we must have the usual embedding constraint

$$
E^\alpha = 0,
$$

(55)

by similar argument to the one given in the preceding discussion of string ending on branes.

### 3.2. An example: the D2-brane ending on a D4-brane

To illustrate the general formalism introduced above we consider an open D2-brane ending on a D4-brane. According to the results of the preceding section there will be two potentials, the usual one-form potential $A_1$, and the dual two-form potential $A_2$, on the D4-brane. The Bianchi identities are (although the mass parameter $m$ does not arise in the Bianchi identities (46) and (53), we will nonetheless set it to zero for simplicity)

$$
d\mathcal{F} = -H,
$$

(56)

and

$$
d\mathcal{F}_3 = G_4 + G_2 \mathcal{F}.
$$

(57)

In addition we are required to take all the components of both $\mathcal{F}$ and $\mathcal{F}_3$ to be zero, except for those which have solely vectorial indices.

In the case of the D4-brane the embedding constraint is enough to force the equations of motion [2,5] which are usually written in terms of the one-form potential $A_1$. However, there is also a dual GS version [13] in which one replaces $A_1$ with a two-form $A_2$. Since in the superembedding formalism the brane is on-shell due to the basic constraint (55) it follows that we should be able to construct either, or indeed both, versions, and the open brane set-up naturally gives both potentials.

To analyse the above Bianchi identities we set

$$
E^\alpha_a = u^a_a + h^a_\beta u^{-\beta}_a,
$$

(58)

and

$$
E^\alpha_a = u^a_a,
$$

(59)

where $u$ denotes an element of $Spin(1,9)$, in either the spin representation or the vector representation according to the indices. The basic constraint (55) implies that

$$
E^\alpha_a E^\beta_a \delta_{\alpha \beta}^\gamma = T^a_a \gamma^a_c E^\gamma_c,
$$

(60)

from which one finds that [5]

$$
h^{a}_{\alpha} \gamma^{\beta} \rightarrow h^{a}_{\alpha} \gamma^{\beta} = i \delta_{\alpha} (\gamma^{ab})_{\alpha}^{\beta} h^{a}_{\beta},
$$

$$
T_{a \beta \gamma} = -i \eta_{\alpha} (\gamma^{ab})_{\alpha \beta} m^{a}_{\gamma} + C_{\alpha \beta} m^{a}_{\gamma}.
$$

(61)
Here we have introduced the two-step notation for the spinor indices on $M$. The index $a$, running from 1 to 16, is rewritten as the pair $a_i$, where $a$ is a five-dimensional spinor index running from 1 to 4, and $i$ is an $Sp(4)$ index, also running from 1 to 4.

The $m$-tensors are given by

$$m_{ab} = (1 - 2y_1)\eta_{ab} + 8(h^2)_{ab},$$

$$m^a = -i\epsilon^{abcd}h_{bc}h_{de},$$

where $y_1$ and $y_2$ denote the two invariants

$$y_1 = tr h^2,$$

$$y_2 = tr h^4.$$  \hspace{1cm} (63)

It is straightforward to check the Bianchi identities for the $F$'s using this information. If we take the target space to be flat for simplicity, the only non-vanishing components of the RR tensors $G_{\alpha\beta}$ and $G_2$ are

$$G_{\alpha\beta\gamma\delta} = -i(\Gamma_{ab})_{\alpha\beta\gamma\delta},$$ \hspace{1cm} (64)

and

$$G_{\alpha\beta} = -i(\Gamma_{11})_{\alpha\beta},$$ \hspace{1cm} (65)

while the non-vanishing component of the NS tensor $H$ is

$$H_{\alpha\beta\gamma} = -i(\Gamma_1\Gamma_{11})_{\alpha\beta\gamma}.$$ \hspace{1cm} (66)

The dimension zero component of the $F$ Bianchi identity (56) is found to be satisfied if

$$m_{\alpha\beta}\mathcal{F}_{\alpha\beta} = 4h_{ab},$$ \hspace{1cm} (67)

which can be rewritten as

$$\mathcal{F}_{ab} = \frac{4}{(1 + 4y_1^2 - 16y_2)} \times ((1 + 2y_1)h_{ab} - 8(h^3)_{ab}).$$ \hspace{1cm} (68)

In obtaining this result, it is useful to note the identity

$$X^3 = \frac{1}{2}X^{1\text{tr}X^2 - \frac{1}{2}X(\text{tr}X)^2} + \frac{1}{2}X\text{tr}X^2,$$

which holds for any $5 \times 5$ matrix $X$. One also finds that the dimension zero component of the $F_3$ Bianchi identity (57) is satisfied if

$$\tilde{\mathcal{F}}_{ab} = \frac{4}{(1 + 4y_1^2 - 16y_2)} \times ((1 - 2y_1)h_{ab} + 8(h^3)_{ab}),$$ \hspace{1cm} (69)

where $\mathcal{F}_{ab}$ is the dual of $\mathcal{F}_{abc}$,

$$\tilde{\mathcal{F}}_{ab} = -\frac{1}{3!}\epsilon_{abcde}\mathcal{F}^{cde}. \hspace{1cm} (70)$$

It is in fact enough to show that the dimension zero components of these identities are satisfied to show that the complete identities are. Furthermore, we know from [2,5] and from the string discussion that the worldvolume multiplet for a D4-brane with a one-form potential satisfying the standard constraints is on-shell. It is easy to confirm that this is still the case here by considering the linearised limit in which it becomes clear that $\mathcal{F}_{abc}$ is the dual of $\mathcal{F}_{ab}$. Since the three-form field strength is not a new independent field the version we have derived here is also on-shell.

4. Comments

In this paper we have shown that the equations describing various branes in superspace can be derived by considering the $\kappa$-symmetry of an open brane of appropriate dimension ending on the brane of interest. It is remarkable that the $\kappa$-symmetry considerations for open superbarens, within the framework of superembedding approach, give rise to dual formulations of $D$-branes automatically. Traditional methods to derive the dual $D$-brane actions have been considered in [14]. Here, we find that the potential dual to the usual Born-Infeld vector is furnished in a natural fashion by a suitable $p$-form that lives on the boundary of an open $D$-brane which supports its own Maxwell field.

The results presented here, in our opinion, also furnish further evidence for the power of superembedding approach to a geometrical and elegant description of all superbranes. Indeed, this approach should be applicable to the description of type IIA/B solitonic fivebranes and type I strings/fivebranes as well. It would also be interesting to consider a limit of the model considered here to extract an action for self-dual string in six dimensions. Yet another possible application would be a description of longer superembedding chains or brane networks. Results in this direction will be reported elsewhere [15].
The formalism described here is a hybrid one involving bosonic worldsurface of the first brane but the superworldsurface of the brane to be investigated. One may envisage an approach in which the open brane worldsurface is also elevated into a superspace. This would make the target space and worldsurface supersymmetry manifest and moreover in this approach the geometrical meaning of $\kappa$-symmetry as odd diffeomorphisms of the superworldsurface would become more transparent. Indeed a purely superspace description of open superbranes ending on other superbranes is possible, as we will be shown elsewhere [15].

**Acknowledgements**

We thank Per Sundell for helpful discussions. One of the authors (E.S.) thanks the Abdus Salam International Center for Theoretical Physics in Trieste for hospitality.

**Appendix A**

Here we collect the properties of the various $\Gamma$ matrices in diverse dimensions.

For type IIA, we use the conjugation matrices,

$$C_{a\beta} = \begin{cases} 
-i\sigma_2 \times C \times \eta, & p = 0,4,8 \\
1 \times C \times \eta, & p = 2,6 
\end{cases} \quad (A.1)$$

and the $\Gamma$-matrices

$$(\Gamma_2)_{\alpha\beta} = \begin{cases} 
\sigma_1 \times \gamma^a \times 1, & p = 0,4,8 \\
\sigma_1 \times 1 \times \gamma^a, & p = 2,6 
\end{cases} \quad (A.2)$$

where $\gamma^a$ and $\gamma^a$ are the $\gamma$-matrices, while $C$ and $\eta$ are the charge conjugation matrices of $SO(p,1)$ and $SO(9-p)$, respectively. The matrices $\gamma^a$ and $\eta$ are symmetric for $p = 0,2,8$ and antisymmetric for $p = 4,6$, while the matrices $C$ and $\gamma^a$ are symmetric for $p = 0,6,8$ and antisymmetric for $p = 2,4$.

The chirality matrix $\Gamma_{11} = \Gamma_0 \Gamma_1 \cdots \Gamma_9$ is given by

$$(\Gamma_{11})_{\alpha\beta} = (\sigma_2 \times 1 \times 1)_{\alpha\beta}. \quad (A.3)$$

A 32 component Majorana spinor $\psi$ in ten dimensions decomposes as

$$\psi^{\underline{2}} = \begin{pmatrix}
\psi^\alpha \\
\psi^{\alpha'}
\end{pmatrix} \quad (A.4)$$

where $\alpha = 1, \ldots, 16$ labes the fermionic coordinates of the worldvolume, and $\alpha' = 1, \ldots, 16$ labels the fermionic transverse directions. The $\sigma$- matrix factors of the $\Gamma$-matrices act on the doublet (A.4). Thus, we have for example,

$$\Gamma_{\alpha\beta}^a = 0, \quad \Gamma_{\alpha'\beta'}^{a'} = -C_{a\beta} \eta_{ij}, \quad \Gamma_{a\beta} = C_{a\beta} \gamma_{ij}^a \quad (A.5)$$

where $i = 1, \ldots, 9 - p$ label the vector representation of the transverse $SO(9-p)$.

In the case of type IIB $\Gamma$-matrices, we suppress the ten dimensional $SO(2)$ doublet index, and consider the chirally projected $16 \times 16 \Gamma$-matrices. The unprimed spinor index labelling the fermionic worldvolume coordinates, and the primed spinor indices labelling the transverse fermionic directions are denoted by using the projection operators

$$P_\pm = \frac{1}{2}(1 \pm \sigma_3), \quad (A.6)$$

acting on the $SO(2)$ indices $I,J = 1,2$, as follows:

$$\psi_\alpha = (P_+ \psi)^\alpha, \quad (A.7)$$

Now we can construct the ten dimensional $\Gamma$-matrices as:

$$p = 9: \quad \Gamma_9^\alpha = \gamma_{\alpha\beta} P_+$$

$$p = 7: \quad \Gamma_7^\alpha = \left(\begin{array}{c}
\sigma_{\alpha\beta} P_+ \\
C_{\alpha\beta} P_-
\end{array}\right)$$

$$p = 5: \quad \Gamma_5^\alpha = \left(\begin{array}{c}
\gamma_{\alpha\beta} \eta_{ij} P_+ \\
\delta_{\alpha\beta} \gamma_{ij}^{a'} P_-
\end{array}\right)$$

$$p = 3: \quad \Gamma_3^\alpha = \left(\begin{array}{c}
\sigma_{\alpha\beta} \delta_i P_+ \\
C_{\alpha\beta} \gamma_{ij}^a P_-
\end{array}\right)$$

$$p = 1: \quad \Gamma_1^\alpha = \left(\begin{array}{c}
\gamma_{\alpha\beta} \delta_i P_+ \\
C_{\alpha\beta} \gamma_{ij}^a P_-
\end{array}\right) \quad (A.8)$$

where $\gamma_{\alpha\beta}$ and $\sigma_{\alpha\beta}$ are the chirally projected $\gamma$-matrices appropriate to ($p+1$)-dimensions. For $p =$
9, there is no transverse direction, and consequently $a = a = 0, 1, \ldots, 9$. In the case of $p = 7$, the two transverse coordinates have been combined into a single complex coordinate. Further details can be found in [5].

References

Fermionic zero mode and string creation between D4-branes at angles

Takuhiro Kitao a,1, Nobuyoshi Ohta b,2, Jian-Ge Zhou b,3

a Institute of Physics, University of Tokyo, Komaba, Tokyo 153, Japan

b Department of Physics, Osaka University, Toyonaka, Osaka 560, Japan

Received 20 January 1998

Editor: M. Dine

Abstract

We study the creation of a fundamental string between D4-branes at angles in string theory. It is shown that $R(-1)^F$ part of the one-loop potential of open string changes its sign due to the change of fermionic zero-mode vacua when the branes cross each other. As a result the effective potential is independent of the angles when supersymmetry is partially unbroken, and leads to a consistent picture that a fundamental string is created in the process. We also discuss the s-rule in the configuration. The same result is obtained from the one-loop potential for the orthogonal D4-branes with non-zero field strength. The result is also confirmed from the tension obtained by deforming the Chern-Simons term on one D4-brane, which is induced by another tilted D4-brane. © 1998 Elsevier Science B.V. All rights reserved.

Recently it has been pointed out from various points of view that a brane (string) is created when certain two branes cross each other [1–14]. This is originally suggested in the field theory analysis on D-branes in Ref. [1]. Consistencies of this result are confirmed by deforming the Chern-Simons term of the system [3] and by the one-loop potentials of open strings [4]. Moreover, the results of M(atrix) theory also support this fact [6,8] and the authors of Refs. [10–12] have explained the brane creation by using the technique of holomorphic embedding of M5-brane or M2-brane into Taub-NUT space induced by Kaluza-Klein monopole.

In this paper, we study the creation of a fundamental string between two D4-branes at angles, in which the configuration has less supercharges than those of orthogonal case [15–22]. This system was examined in the framework of the M(atrix) theory in Ref. [8], and we intend to study it and its variant in string theory in more detail. Of course, we expect that the result is the same as that of M(atrix) theory [8] because the approach of M(atrix) theory necessarily involves the theory of bound states of D-branes whose T-duality is the theory of branes with angles discussed in Ref. [20]. However, we find that there is a subtle issue which needs clarification.
To discuss the string creation, we use the one-loop potential due to the open strings between the branes at angles. Amplitudes for such system have been obtained in Ref. [21] to determine the condition of unbroken supersymmetry, and we can simply read off the potential from that work. We also obtain the one-loop potential from open strings between the orthogonal branes with condensation which is similar to the case with angles. When supersymmetry is partially unbroken, these potentials vanish. The problem is what happens when one brane adiabatically crosses another brane, keeping the partially unbroken supersymmetry. The authors of Ref. [4] have interpreted this situation as the system of brane-anti-brane, and obtained the result which is consistent with the creation of a fundamental string, that is, the effective potential of the string tension times the distance between the branes, which is canceled by a string created between these branes. They discussed the case that D0-brane passes through D8-brane, in which the NS$(-1)^F$ term vanishes because of the fermionic zero modes.

On the other hand, in the case with angles, NS$(-1)^F$ term gives generally non-zero contribution which depends on angles. As a consequence, it turns out that if we interpreted the system as that of brane-anti-brane when one brane adiabatically crossed another, the potential would be dependent on the angles. This is a very strange result if we try to understand it in terms of a string stretched between the two branes. We find that in our present situation only the R$(-1)^F$ term changes the sign after brane crossing. Our results also indicate that the s-rule of Hanany and Witten is valid [1]. The rule translated into our present setting states that a configuration with more than one fundamental string joining the two D4-branes cannot be supersymmetric. As we will see, this rule can be understood as the uniqueness of the chiral fermionic zero-mode vacuum.

Let us start with the configuration of two D4-branes, one of which has tilted world-volume to that of another D4-brane. The boundary conditions of the open string on one D4-brane (denoted as D4) are

\[
\delta X^\mu = \bar{\psi} \gamma^\mu.
\]

Those at \(\sigma = \pi\) on another D4-brane (denoted as D4') are

\[
\delta X^0 = 0,
\]

\[
\delta \sigma X^i \cos(\theta, \pi) - \delta \sigma X^{i+4} \sin(\theta, \pi) = 0,
\]

\[
X^i \sin(\theta, \pi) + X^{i+4} \cos(\theta, \pi) = 0, \quad i = 1, \ldots, 4,
\]

\[
X^0 = b,
\]

where \(b\) is the distance between the two branes and \((\theta, \pi)\) are the angles parameterizing our two D4-brane configuration.

The one-loop amplitude of the open string which satisfies the above boundary conditions has been obtained in Ref. [21]. Assuming that all the four
angles are non-zero, \(^4\) we find that the potential between these D4-branes is
\[
V = \left. - \int_0^\infty \frac{dt}{2t} \frac{e^{-\frac{b^2t}{(2\pi\alpha')^2}}}{(8\pi^2\alpha')^2} \right|^{1\over 2} 
\times \left[ - \frac{4}{k=1} \Theta_j(i\theta_j t|\bar{t}|) + \frac{4}{k=1} \Theta_k(i\theta_k t|\bar{t}|) \right. 
\left. + \frac{4}{k=1} \Theta_l(i\theta_l t|\bar{t}|) \right] 
\right|^{1\over 2}.
\tag{4}
\]

The terms in the bracket are the contributions from R, NS, R(−1)^K and NS(−1)^K sectors, respectively. The author of Ref. [21] has derived the condition that the amplitude vanishes, in the search for the criterion that supersymmetry is partially unbroken. It is given by
\[
\theta_1 \pm \theta_2 \pm \theta_3 \pm \theta_4 = 0, \mod 2,
\tag{5}
\]
for which we have
\[
\frac{4}{k=1} \Theta_j(i\theta_j t|\bar{t}|) - \frac{4}{k=1} \Theta_k(i\theta_k t|\bar{t}|) 
+ \frac{4}{k=1} \Theta_l(i\theta_l t|\bar{t}|) - \frac{4}{k=1} \Theta_k(i\theta_k t|\bar{t}|) = 0.
\tag{6}
\]

Indeed, Eq. (5) is the same as that derived in supergravity in Ref. [22], in which it has been shown that there is 1/16 unbroken supersymmetry if Eq. (5) is satisfied. In what follows, we will mainly concentrate on the case (5).

Let us consider what happens when the above one brane passes through another adiabatically in the direction of \(x^9\) from \(x^9 = |b|\) to \(x^9 = -|b|\). We can regard this situation as the sum of two parts. One is that of one of the branes hooked by another brane, which is expected to be a fundamental string. Another part is that from the two branes passing by without hooking. If the interaction between the two branes of the latter part in this situation is the string tension times the distance between the branes, the first part is expected to represent the created fundamental string to cancel this interaction between the two branes.

What is the difference between the state before crossing (‘‘configuration [A]’’) in which 1/16 SUSY is unbroken and that after crossing (‘‘configuration [B]’’) ignoring the hooking part? The only difference is the vacua defined by the R-fermion zero modes, which exist only in the world-sheet fermions \(\psi^0\) and \(\psi^3\). There are no other zero modes because the boundary conditions with the angles shift the R-fermion zero modes of the rotated directions. We define the vacua \(|\pm\rangle\) by
\[
(\psi^0_1 \mp \psi^0_0)|\pm\rangle = 0.
\tag{7}
\]

By GSO projection, either + or − is projected out, so that only one space-time massless chiral fermion can exist when the two branes intersect. Configuration [A] is related to [B] by the parity transformation in the direction of \(x^9\). In order to preserve the supersymmetry (2), \(\psi^0_0\) transforms into \(-\psi^0_0\) when the configuration changes from [A] to [B]. It follows that the vacuum of [A] is different from that of [B]. In other words, the definition of GSO projection for the R-sector is different between [A] and [B]. As a result, the potential of [B] is different from [A] by the sign of the R(−1)^K term. By using the condition of unbroken supersymmetry (6), we find
\[
V = -\frac{1}{2} \int_0^\infty \frac{dt}{t} \frac{e^{-\frac{b^2t}{(2\pi\alpha')^2}}}{(8\pi^2\alpha')^2} \frac{1}{t} 
\times \left[ - \frac{4}{k=1} \Theta_j(i\theta_j t|\bar{t}|) + \frac{4}{k=1} \Theta_k(i\theta_k t|\bar{t}|) 
\pm \frac{4}{k=1} \Theta_l(i\theta_l t|\bar{t}|) - \frac{4}{k=1} \Theta_k(i\theta_k t|\bar{t}|) \right] 
\right|^{1\over 2}.
\tag{8}
\]

\(^4\) For zero angle \(\theta_i\), it is easy to do a similar calculation. One finds that the summation of the bosonic degrees of freedom is replaced by the product of the integral over the momentum \(p_i\) and the volume of the direction \(X^i\). It should be possible to continuously interpolate the zero angle limit in Eq. (4), but it seems complicated to do so explicitly.

\(^5\) More concrete picture of this ‘‘hooking’’ will be explained below.
where \((- (+)\) in the last equation is for the configuration \([A] ([B])\), and \(T_0\) is the string tension \(\frac{1}{(2\pi\alpha')}\). This change of the sign occurs when the two branes intersect each other at \(b = 0\). We see that the potential computed in Eq. (8) for \([B]\) is independent of the angles and equals the string tension times the distance. On the other hand, since we move one brane across another from \(x^0 = |b|\) to \(x^0 = -|b|\) adiabatically, the fermionic zero-mode vacuum is changed after crossing from \(x^0 = 0\) all the way down to \(x^0 = -|b|\) and we get contributions from that part which exactly cancel the above potential. This is what we have called “hooking” in the above. Physically this should be understood as due to a string created between the two branes when supersymmetry is unbroken. In this way, the BPS property of the system is preserved before and after the brane crossing.

It is crucial that only the \(R(-1)^F\) term changes its sign; if we interpreted the final configuration as the special orthogonal case can be transformed to the RR sector in the closed string channel would change the signs, resulting in a potential dependent on the angles. Such potential could not be canceled by string tension, and would allow no physical interpretation.\(^6\)

The fact that there is only one space-time massless chiral fermion exists at the intersection is closely related to the \(s\)-rule \([1, 12, 14]\). Our results indicate that only a single string is stretched between two D4-branes after crossing. By a chain of dualities, the special orthogonal case can be transformed to the configuration in which a D3-brane is suspended between an NS 5-brane and a D5-brane considered in Ref. [1]. Thus our results give the generalization of this rule to more general angles.

It is easy to repeat the calculation when any one of the angles \(\theta_i\) is zero under the condition (5), and we find that the potential vanishes and no string is created.

We can derive the same result in the case of orthogonal D4-D4' system with bound state \([24–29]\).

The boundary conditions of the open string on one D4-brane are

\[
\begin{align*}
\partial_\sigma X^0 &= 0, \\
\partial_\sigma X^{2k-1} + 2\pi\alpha' F_{(k)} &\partial_\sigma X^{2k} = 0, \\
\partial_\sigma X^{2k} - 2\pi\alpha' F_{(k)} &\partial_\sigma X^{2k-1} = 0, \quad k = 1, 2, \\
X^\mu &= 0, \quad \mu = 5, \ldots, 9,
\end{align*}
\]

at \(\sigma = 0\), where \(\tau\) is the world-sheet coordinate along with \(\sigma\) and \(F_{(k)}\) is the condensation of the field strength. The boundary conditions at \(\sigma = \pi\) on another D4'-brane are

\[
\begin{align*}
\partial_\tau X^0 &= 0, \\
X^\mu &= 0, \quad \mu = 1, \ldots, 4, \\
\partial_\tau X^{2k-1} + 2\pi\alpha' F_{(k)} &\partial_\tau X^{2k} = 0, \\
\partial_\tau X^{2k} - 2\pi\alpha' F_{(k)} &\partial_\tau X^{2k-1} = 0, \quad k = 3, 4, \\
X^b &= b.
\end{align*}
\]

By repeating the calculation similar to the above, we obtain the potential

\[
V = -\int_0^\infty dt \frac{e^{-x^2/2}}{(2\pi\alpha')^{\frac{3}{2}}} \times \left[ -\frac{4}{\Theta_3(i\epsilon_k t|it)} + \frac{4}{\Theta_3(i\epsilon_k t|it)} - \frac{4}{\Theta_3(i\epsilon_k t|it)} \right],
\]

where

\[
\epsilon_k = \frac{\tan^{-1}(2\pi\alpha' F_{(k)})}{\pi}
\]

and \(\pm\) corresponds to the difference in the definitions of GSO projection. The condition for \(1/16\) unbroken SUSY is the same as the case with angles:

\[
\epsilon_1 \pm \epsilon_2 \pm \epsilon_3 \pm \epsilon_4 = 0, \quad \text{mod} \ 2.
\]
With the help of Eq. (6), the potential (11) is cast into
\[
V = -\int_0^{\infty} \frac{dt}{2t} \frac{e^{-\frac{b^2 t}{(2\pi t)^{\frac{3}{2}}}}}{(8\pi^2 \alpha')^2} (-1 \pm 1)
\]
\[
= -\frac{\mu_{D4}}{2} (1 \mp 1) |b|.
\]
(13)

This also shows that a fundamental string is created.

As a further check of our results, let us calculate the induced tension from the Chern-Simons term. Following Ref. [3] in which the system of D0 strings is discussed, we consider the Chern-Simons term on the first D4:
\[
\mu_{D4} \frac{1}{4!} \int d^5 x \epsilon_{x_0 \ldots x_4} F_{(4)}^{x_0 \ldots x_4} A^{x_4},
\]
(14)

where \( F_{(4)} \equiv dC_{(3)} \) is the R-R 3-form gauge field, and \( A^x \) is the U(1) gauge field on D4. The indices \( \{ x_i \} \) are those of the coordinates of the world-volume of the first D4, \( \ldots \), \( A^x \) depends on only \( x_i \) and \( A^x \) depends on only \( x_i \). The above term reduces to
\[
\mu_{D4} \int \left( \prod_{i=1}^{4} dx_i \right) F_{(4)}^{1234} \int ds \frac{dx_0}{ds} A^0(x_0),
\]
(15)

where \( s \) is a parameter on the D4 and can be supposed to be a world-sheet coordinate, and then \( \mu_{D4} / (\prod_{i=1}^{4} dx_i) F_{(4)}^{1234} \) is interpreted as the tension to be evaluated below.

On the other hand, the R-R charge \( \mu_{D3} \) of D4 is
\[
\mu_{D3} = \int_{S_4} * F_{(4)} d^4 x = \int_{S_4} F_{(4)} d^3 x,
\]
(16)

where \( S_4 \) is the 4-sphere surrounding D4. Hence \( F_{(4)} \) can be taken as
\[
\frac{1}{4!} \epsilon_{x_0 x_1 x_2 x_3 x_4} F_{(4)}^{x_0 x_1 x_2 x_3} = \mu_{D4} x_4 \Omega_4 / r,
\]
(17)

where all \( x_i \) are the indices of the directions which are orthogonal to the world-volume of D4, and \( r \) and \( \Omega_4 \) are the radius of \( S_4 \) and the volume of a unit 4-sphere, respectively. We denote the directions of the world-volume of D'4 as \( y_0, \ldots, y_4 \). Then the non-zero components of \( F_{(4)}^{y_0 \ldots y_4} \) are those with indices outside these dimensions:
\[
F_{(4)}^{1234}, F_{(4)}^{1235}, F_{(4)}^{1245}, F_{(4)}^{1345}, F_{(4)}^{2345}, F_{(4)}^{12345}.
\]
(18)

To evaluate the contribution to \( F_{(4)}^{1234} \) at a point Q on D4 from a point P on the second D4, let us consider the 5-dimensional plane which intersects D4 orthogonally at P, and also D4 at Q. The relation between the coordinates \( \{ x_i \} \) for D4 and \( \{ x_i \} \) for D'4 is
\[
\begin{align*}
x_i &= x_i' \cos(\theta_i \pi) + x_{i+4} \sin(\theta_i \pi), \\
x_{i+4} &= x_{i+4}' \cos(\theta_i \pi) - x_i' \sin(\theta_i \pi), \quad i = 1, \ldots, 4, \\
x_9 &= x_9' + b.
\end{align*}
\]
(19)

We see that the coordinates \( x_1, \ldots, x_8 \) mix with \( x_1, \ldots, x_4 \) with the coefficients \( \sin(\theta_i \pi) \) and hence \( F_{(4)}^{1234} \) in (15) gets the following contribution from \( F_{(4)}^{1234} \):
\[
F_{(4)}^{1234} = \left( \prod_{i=1}^{4} \sin(\theta_i \pi) \right) F_{(4)}^{1234} = -\frac{\mu_{D3}}{r \Omega_4} \frac{b}{r} \left( \prod_{i=1}^{4} \sin(\theta_i \pi) \right),
\]
(20)

where \( r \) is the distance between \( P \) and \( Q \), and we have used Eq. (17) and the fact that \( P \) is away from \( Q \) by \( b \) in the direction of \( x_9 \). Let us take the origin on D4 at the nearest point from D'4 and denote the coordinates of Q as \( (y_1, y_2, y_3, y_4) \). Then \( r = PQ \) is given by \( r^2 = b^2 + \sum_{i=1}^{4} y_i^2 \sin^2(\theta_i \pi) \). Using this expression and Eq. (20), we obtain
\[
\begin{align*}
\mu_{D3} &\int \left( \prod_{i=1}^{4} dy_i \right) F_{(4)}^{1234} = \mu_{D3} \int \left( \prod_{i=1}^{4} dy_i \right) \left( \prod_{i=1}^{4} \sin(\theta_i \pi) \right) F_{(4)}^{1234} \\
&= -\mu_{D3} \int \left( \prod_{i=1}^{4} dy_i \right) \left( \prod_{i=1}^{4} \sin(\theta_i \pi) \right) \frac{\mu_{D3} b}{r \Omega_4} \left( \prod_{i=1}^{4} \sin(\theta_i \pi) \right) \\
&= -\frac{1}{2} \mu_{D3} \mu_{D3} \frac{b}{|b|},
\end{align*}
\]
(21)
which is independent of the angles. Because

\[ \mu_{\text{D4}} = \frac{1}{2 \pi \alpha'} \left( \frac{1}{2} \right), \]

Eq. (21) turns out to be

\[ \frac{1}{4 \pi \alpha'} \frac{b}{|b|} = - \frac{T_0}{2} \frac{b}{|b|}, \]

that is, one half contribution of the interaction between a fundamental string and a D-brane. \(^7\) This result is the same as the orthogonal case of [3].

Let us compare this result with our previous discussions on the potential. Consider the configuration [A]. The above tension is the same as the force from the \( R(-1)^F \) term in (4) up to sign even if (5) is not satisfied. This means that the anomaly term corresponds to that of the created string. The double of the force induced was computed in Eq. 8 and the second term is for the lightest mass in supergravity. In fact the authors of [8] have derived the same term by doing so in M(atrix) theory.

To consider the configuration [B], we divide the one-loop potential of configuration [A] into two pieces as

\[ R + \text{NS} + \text{NS}(-1)^F + R(-1)^F \]

\[ = (R + \text{NS} + \text{NS}(-1)^F - R(-1)^F) \]

\[ + 2R(-1)^F. \quad (22) \]

The left hand side is the potential of configuration [A] which vanishes by supersymmetry. The first term on the right hand side is the potential for [B] which we computed in Eq. (8) and the second term is for the created string. The double of the force induced by the anomaly term corresponds to that of the created string, which gives the same picture as in Ref. [3].

String creation in type I' theory has also been discussed in Refs. [4,9,12]. It should be straightforward to generalize our above discussions to similar configurations in such theory, and this will not be discussed in this paper.

In summary, we have derived the results which are consistent with the creation of a fundamental string in the crossing process of D4-branes at angles and its variant configuration by using the potentials of string and deformation of the anomaly term. It is consistent with the results of M(atrix) theory [8]. In the process, we have clarified how the configurations, and in particular, the fermionic zero-mode vacua change, giving a consistent picture of string creation. Our results also confirm the s-rule discussed for orthogonal case.

T.K. would like to thank the people of Particle Theory Group at Komaba for helpful comments, especially T. Yoneya for useful discussions. The work of N.O. and J.-G.Z. was supported in part by Grand-in-aid from the Ministry of Education, Science, Sports and Culture No. 96208.

References


---

\(^7\) The overall sign is not significant.

\(^8\) When one of the four angles is zero, both terms vanish, irrespectively of whether supersymmetry is unbroken or not.
A note on the supersymmetries of the self-dual supermembrane

E.G. Floratos a,b, G.K. Leontaris c

a NRCS Demokritos, Athens, Greece
b Physics Department, University of Crete, Iraklion, Crete, Greece
c Theoretical Physics Division, Ioannina University, GR-45110 Ioannina, Greece

Received 4 February 1998
Editor: L. Alvarez-Gaume

Abstract

In this letter we discuss the supersymmetry issue of the self-dual supermembranes in $(8 + 1)$ and $(4 + 1)$-dimensions. We find that all genuine solutions of the $(8 + 1)$-dimensional supermembrane, based on the exceptional group $G_2$, preserve one of the sixteen supersymmetries while all solutions in $(4 + 1)$-dimensions preserve eight of them. © 1998 Elsevier Science B.V. All rights reserved.
We start by recalling the light-cone gauge formulation of the supermembrane, where half of the rigid space-time supersymmetry as well as the local $\kappa$-symmetry is fixed. We then provide the supersymmetries left intact.

It is known that in $d = 8$ dimensions there is a connection of the Clifford algebra with the octonionic algebra and this is the information needed to study the behaviour of the octonionic self-duality equations under supersymmetry transformations. The relation with the octonions has been noticed in the 80’s during the studies of the $S^7$-compactifications of the 11-d supergravity as well as for the $N = 8$ gauged supergravities [8–10]. Recently, the embedding of octonionic Yang-Mills (YM) instantons in the ten-dimensional effective supergravity theories of strings has been constructed [11–13] where it was found that one supersymmetry survives. More generally, wrapped membrane compactifications have been recently discussed in the literature [14].

In the light-cone gauge, after the elimination of the $X_-$ variable from the reparametrization constraints, the supersymmetric Hamiltonian [6,15] is defined as

$$\mathcal{H} = \frac{1}{P_0^2} \int d^3x \left( \frac{1}{2} P^I P_I + \frac{1}{2} \{ X^I, X^J \}^2 - P_0^+ \bar{\theta} \Gamma_- \{ X^I, \theta \} \right)$$

(1)

where $P_I = \dot{X}_I$ and the indices $I,J = 1,\ldots,9$ while we have fixed the area preserving parameters so that $w = 1$ [6]. The compatibility condition for the uniqueness of $X_-$, is the Gauss law

$$\{ \dot{X}_I, X_J \} + \{ \bar{\theta} \Gamma_- \theta \} = 0, \ I = 1,\ldots,9$$

(2)

where summation over repeated indices is assumed. The Clifford generators $\Gamma_I$, in (1) are represented by real 32 $\times$ 32 matrices which can be chosen in the following form

$$\Gamma_I = \sigma_3 \otimes \gamma_I$$

(3)

where $\sigma_3$ is the Pauli matrix, $\gamma_I$ represent the 16 $\times$ 16 matrices and $\gamma_0 = \gamma_1 \cdots \gamma_9$. Further, $\Gamma_-$ (and $\Gamma_+$) correspond to the light-cone coordinates $(X_{10} \pm X_0)/\sqrt{2}$, thus they are given by a similar decomposition

$$\Gamma_\pm = \frac{1}{\sqrt{2}} (\Gamma_{10} \pm \Gamma_0).$$

(4)

Thus, we have

$$\Gamma_- = i\sqrt{2} \left( \begin{array}{cc} 0 & 1_{16} \\ 0 & 0 \end{array} \right), \quad \Gamma_+ = i\sqrt{2} \left( \begin{array}{cc} 0 & 0 \\ 1_{16} & 0 \end{array} \right).$$

(5)

The Hamiltonian (1) is invariant under area-preserving transformations of the membrane (which for non-trivial topologies of the membrane contain also global elements $2g$ in number, where $g$ is the genus of the membrane [15]). The local area-preserving transformations are generated by the Gauss law (2). Here the canonical variables satisfy Dirac brackets

$$\{ X^I(\sigma), \dot{X}^J(\sigma') \}_{DB} = \delta^{IJ} \delta(\sigma - \sigma')$$

(6)

$$\{ \dot{\sigma}^I(\sigma), \dot{\sigma}^J(\sigma') \}_{DB} = \frac{i}{4} (\Gamma_+)^{IJ} \delta^2(\sigma - \sigma')$$

(7)

(where we have chosen $P_+ = 1$). It can be verified that in the light-cone gauge there are two independent spinor supersymmetry charges

$$Q = Q^+ + Q^- = \int d^2 \dot{\sigma} J^0$$

(8)
where \( Q^\pm = \frac{1}{2} \Gamma_+ Q \), and

\[
Q^+ = \int d^2 \sigma \left( 2 \dot{X}^i \Gamma^i_j + \{ X^i, X^j \} \Gamma^i_\sigma \right) \theta
\]

\[
Q^- = 2 \int d^2 \sigma \delta = 2 \Gamma_+ \theta_0
\]

which are constants of motion and \( \theta_0 \) is the momentum conjugate to the center-of-mass coordinate of the fermionic degrees of freedom of the membrane. The corresponding supersymmetry transformations which leave the Hamiltonian invariant, are given by

\[
\delta X^i = -2 \varepsilon \Gamma^i \theta
\]

\[
\delta \theta = \frac{1}{2} \Gamma_+ \left( \dot{X} \Gamma^i_j + \Gamma_- \right) \epsilon + \frac{1}{2} \{ X^i, X^j \} \Gamma_+ \Gamma^j \delta \epsilon.
\]

On the other hand, the local fermionic \( \kappa \)-symmetry has been fixed by imposing the condition

\[
\Gamma_+ \theta = 0.
\]

Due to this gauge condition, the fermionic coordinates are restricted to \( SO(9) \) spinors, satisfying

\[
\Gamma_1 \cdots \Gamma_9 \theta = \theta
\]

while the \( SO(9) \) \( \Gamma \)-matrices satisfy \( \Gamma^T = \Gamma \).

The self-duality equations for the bosonic part of the supermembrane have been initially introduced in the light-cone gauge fixing \( X_1, \ldots, X_9 \) to be constants [1],

\[
\dot{X}_i = \frac{1}{2} \epsilon_{ijk} \{ X_j, X_k \}, \quad i,j = 1,2,3.
\]

These equations have been proposed as an analogue of the electric-magnetic duality where the local velocity of the membrane corresponds to the electric field while the RHS which is the normal to the membrane surface, corresponds to the magnetic field. They imply the Gauss law and the Euclidean-time equations of motion with fermionic degrees of freedom (dot) set to zero [1,16]. This system has been shown to be integrable and a Lax pair was found. In order to go to higher dimensions one should have the notion of cross product of two vectors and this is provided as the unique other possibility by the structure constants of the algebra of octonions (Cayley numbers) [17]. The octonionic units \( o_i \) satisfy the algebra

\[
o_i o_j = -\delta_{ij} + \Psi_{ijk} o_k.
\]

where \( i = 1, \ldots, 7 \) are the 7 octonionic imaginary units with the property

\[
\{ o_i, o_j \} = -2 \delta_{ij}.
\]

The totally antisymmetric symbol \( \Psi_{ijk} \) appearing in (16) is defined to be equal to 1 when the indices are [17]

\[
\Psi_{ijk} = \begin{pmatrix}
1 & 2 & 4 & 3 & 6 & 5 & 7 \\
2 & 4 & 3 & 6 & 5 & 7 & 1 \\
3 & 6 & 5 & 7 & 1 & 2 & 4
\end{pmatrix}
\]

and zero for all other cases. With this multiplication table, \( \Psi_{ijk} \) provides for every two seven-dimensional vectors a third one, normal to the first two. Thus, it is possible to extend the three-dimensional self-duality equations to seven dimensions, fixing only the values of \( X_8, X_9 \) membrane coordinates. Then, the self-duality equations [2] become

\[
\dot{X}_i = \frac{1}{2} \Psi_{ijk} \{ X_j, X_k \}.
\]
The Gauss law results automatically by making use of the $\Psi_{ijk}$ cyclic symmetry
\[ \{ \hat{X}_r, X_i \} = 0. \] (20)

The Euclidean equations of motion are obtained easily from (19)
\[ \hat{X}_i = \{ X_i, \{ X_i, X_i \} \} \] (21)
where use has been made of the identity
\[ \Psi_{ijk} \Psi_{lmn} = \delta_i \delta_j \delta_k - \delta_i \delta_m \delta_j + \delta_j \delta_i \delta_m \] (22)
and of the cyclic property of the symbol $\phi_{ijkl}$ [17] which is defined to be equal to 1 when its indices take values of the following table:
\[
\begin{array}{cccccc}
4 & 3 & 6 & 5 & 7 & 1 \\
5 & 7 & 1 & 2 & 4 & 3 \\
6 & 5 & 7 & 1 & 2 & 4 \\
7 & 1 & 2 & 4 & 3 & 6 \\
\end{array}
\] (23)
whilst it is zero for any other combination of indices. In terms of these units an octonion can be written as follows:
\[ X = x_0 o_0 + \sum_{i=1}^{7} x_i o_i \] (24)
with $o_0$ the identity element. The conjugate is
\[ \bar{X} = x_0 o_0 - \sum_{i=1}^{7} x_i o_i. \] (25)

The octonions over the real numbers can also be defined as pairs of quaternions
\[ X = (x_1, x_2) \] (26)
where $x_1 = x_1^1 o_1^1$, $x_2 = x_1^2 o_2^2$ and the indices $\mu$ run from 0 to 3, while $x_1^0$ are real numbers and $x_1^i$ ($i = 1, 2, 3$) are imaginary numbers. Finally, $o_0$ is the identity $2 \times 2$ matrix and $o_i$ are the three standard Pauli matrices.

If $q = (q_1, q_2)$ and $r = (r_1, r_2)$ are two octonions, the multiplication law is defined as
\[ q * r = (q_1, q_2) * (r_1, r_2) = (q_1 r_1 - r_2 q_2, r_2 q_1 + q_2 r_1) \] (27)
where $q_1 = q_1^0 + q_1^i o_i$ and $\bar{q}_1 = q_1^0 - q_1^i o_i$. One can also define a conjugate operation for an octonion as
\[ \bar{q} = (q_1, q_2) = (\bar{q}_1, -q_2) \] (28)
and we get the possibility to define the norm and the scalar product $q$ and $r$
\[ q\bar{q} = (q_1 \bar{q}_1 + q_2 \bar{q}_2, 0) = \sum_{\mu=0}^{3} (q_\mu r_\mu + q_\mu \bar{r}_\mu) \] (29)
\[ \langle q|r \rangle = \frac{1}{2} (q \bar{r} + \bar{q} r) \] (30)
In terms of the above formalism, the self-duality equations can be written as follows:
\[ \hat{X} = \frac{1}{2} \{ X, X \} \] (31)
where $X = X^i o_i$ with $i = 1, \cdots, 7$ and the Poisson bracket for two octonions is defined as
\[ \{ X, Y \} = \frac{\partial X}{\partial \xi_1} \frac{\partial Y}{\partial \xi_2} - \frac{\partial X}{\partial \xi_2} \frac{\partial Y}{\partial \xi_1}. \] (32)
After these preliminaries, we come now to the question regarding the number of supersymmetries preserved by the self-duality equations. In our analysis we will explore the number of supersymmetries preserved by (3 + 1)- and (7 + 1)-dimensional solutions. We will see that 3-d solutions preserve as many as eight out of the sixteen supersymmetries while the 7-d self-duality equations preserve only one supersymmetry. The supersymmetry transformation is defined [6]

$$\delta \theta = \frac{1}{2} \left( \Gamma_+ \left( \Gamma_+ \dot{\chi}' + \Gamma_- \right) + \frac{1}{2} \Gamma_+ \Gamma^{ij \ell} \{ X_i, X_j \} \right) \left( i \epsilon_A \epsilon_B \right).$$

In terms of the $16 \times 16$ $\gamma$-matrices, the above is written

$$\delta \theta = \left( \begin{array}{cc} 0 & 0 \\ \sqrt{2} \left( \gamma' \dot{X}_i + \frac{1}{2} \gamma' \gamma^{ij} \{ X_i, X_j \} \right) & -2 \cdot 1_{16} \end{array} \right) \left( \begin{array}{cc} i \epsilon_A \\ \epsilon_B \end{array} \right)$$

which implies that

$$\sqrt{2} \left( \gamma' \dot{X}_i + \frac{1}{2} \gamma' \gamma^{ij} \{ X_i, X_j \} \right) \epsilon_A + 2 \cdot 1_{16} \epsilon_B = 0$$

where $\epsilon_A, \epsilon_B$ are 16-dimensional spinors. From the form of Eq. (34), we observe that if self-duality equations are going to play a role in the preservation of a number of supersymmetries, we should necessarily impose the condition $\epsilon_B = 0$. Thus, at least half of the supersymmetries are broken. Now, the last term in (34) is zero and Eq. (34) simply becomes

$$\left( \gamma' \dot{X}_i + \frac{1}{2} \gamma' \gamma^{ij} \{ X_i, X_j \} \right) \epsilon_A = 0.$$ 

Under the assumption that $\dot{X}_{9,9,9} = 0$, it can be shown that the above reduces to a simpler $8 \times 8$ matrix equation. In order to find a convenient explicit form, we first express the $16 \times 16$ matrices in terms of the octonionic structure constants $C_{ijk}$ as follows: let the index $n$ run from 1 to 7; then we define

$$\gamma_8 = \begin{pmatrix} 0 & 1_s \\ -1_s & 0 \end{pmatrix}, \quad \gamma_n = \begin{pmatrix} 0 & \beta_n \\ -\beta_n & 0 \end{pmatrix}$$

where $1_s$ is the $8 \times 8$-identity matrix and $\beta_n$ are seven $8 \times 8$ $\gamma$-matrices with elements [10]

$$\left( \beta_n \right)_s = \psi_{in}, \quad \left( \beta_n \right)_s = \delta^i_j, \quad \left( \beta_n \right)_s = -\delta^i_j$$

while it can be easily checked that $\beta_1 \cdots \beta_7 = -1_s$ and

$$\gamma_9 = \begin{pmatrix} 1_s & 0 \\ 0 & -1_s \end{pmatrix}.$$ 

The commutation relations of $\beta_n$ give:

$$\left( [ \beta_m, \beta_n ] \right)_s = 2 \psi_{mn},$$

$$\left( [ \beta_m, \beta_n ] \right)_s = -2 \psi_{mn},$$

$$\left( [ \beta_m, \beta_n ] \right)_s = -2 \xi^{mn},$$

where the tensors $\xi^{mn}(u)$ are defined as follows [3]

$$\xi^{ij}(u) = \Delta^{ij} + \frac{u}{4} \phi^{ij}(u)$$

where $\Delta^{ij} = \frac{1}{2} (\delta_i^j \delta^k_l - \delta_i^k \delta^j_l)$. Next, we impose the following condition on the components of the 16-spinor $\epsilon_A$

$$\epsilon_A = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \epsilon.$$
where $\otimes$ stands for the direct product and $\epsilon$ is an eight-component spinor whose components are left unspecified. Clearly, condition (43) reduces further the sixteen supersymmetry charges to eight. Separating the eight components of $\epsilon = (\epsilon_7, \epsilon_1)$ where $\epsilon_{7(\cdot)}$ is a seven-(one-) dimensional vector, we find that Eq. (35) reduces to the matrix equation

$$
\begin{pmatrix}
\Psi_{\mu j} \dot{X}_\mu + \frac{1}{2} \mathcal{H}_{\mu j} (\mathcal{P}^\mu \langle X_\mu, X_\nu \rangle - 4 \langle X_\mu, X_\nu \rangle) \\
0
\end{pmatrix} = 0.
$$

(44)

The rather interesting fact here is that the matrix elements $\sigma_{\epsilon j}$ and $\sigma_{\epsilon k}$, $(j = 1, \ldots, 7)$ multiplying the $\epsilon_1$-component are the self-duality Eqs. (15) in eight dimensions when the Euclidean time-parameter $t$ is replaced with $\mathcal{U}$ (Minkowski). Thus, $\epsilon_1$-component remains unspecified and there is always one supersymmetry unbroken for any eight-dimensional solution of the self-duality equations.

Let us now turn our discussion to the upper $7 \times 7$ part of the matrix Eq. (44). In general, the quantity specifying these elements, namely

$$
\Psi_{\mu j} \dot{X}_\mu + \frac{1}{2} \mathcal{H}_{\mu j} (\mathcal{P}^\mu \langle X_\mu, X_\nu \rangle - 4 \langle X_\mu, X_\nu \rangle)
$$

is not automatically zero. However, there is a particular case –which turns out to be the most interesting one– where the above quantity is the self-duality equation itself. In fact, if we consider only three-dimensional solutions of the equations, the ‘curvature’ factor $\mathbf{f}_{ij}$ is automatically zero while the tensor $\mathcal{H}_{ij}$ simply becomes

$$
\mathcal{H}_{ij} = \Delta_{ij} = \frac{1}{2} \left( \delta_i^k \delta_j^l - \delta_i^l \delta_j^k \right) \quad \text{for} \quad \mathbf{f}_{ij} = 0.
$$

(46)

In this case, it can be easily seen that (45) reduces to the self-duality equations in three-dimensions. In this latter case, all eight supersymmetries survive.

We summarize this note discussing also the importance of the supersymmetric self-duality configurations in three and seven dimensions. The absence of a natural perturbative expansion for the 11-d fundamental supermembrane prohibits so far the derivation of its low energy effective Lagrangian which is expected to contain 11-d, $N = 1$ supergravity interacting with solitonic two- and five-branes in a duality symmetric way. The Euclidean self-dual membrane configurations in three and seven dimensions, after light-cone gauge fixing, provide non-perturbative minima of the action, which could survive perturbative corrections if enough supersymmetries are left intact. Then, the quantum mechanical amplitudes calculated in supermembrane theory could be determined by transforming the path-integral integration around these minima into the infinite moduli-space integration of the self-dual configurations of supermembranes. The best candidate for these seem to be the three-dimensional integrable self-dual sector where eight supersymmetries survive. The problem then is reduced to find the moduli space and its integration measure of the minimum action 3-d configurations. We hope to come back to this problem in a future work.

References

Matrix string interactions

Vipul Periwal, Øyvind Tafjord

Department of Physics, Princeton University, Princeton, NJ 08544, USA

Received 16 February 1998; revised 27 February 1998
Editor: M. Dine

Abstract

String configurations have been identified in compactified Matrix theory at vanishing string coupling. We show qualitatively how the interactions of these strings are determined by the Yang-Mills gauge field on the worldsheet. At finite string coupling, this suggests the underlying dynamics is not well-approximated as a theory of strings. This may explain why string perturbation theory diverges badly, while Matrix string perturbation theory presumably has a perturbative expansion with properties similar to the strong coupling expansion of 2d Yang-Mills theory. © 1998 Published by Elsevier Science B.V. All rights reserved.

String perturbation theory diverges. This divergence has been linked by Shenker [1] to the unusual strength of non-perturbative effects in string theory. Witten [2] has pointed out that the existence of Ramond-Ramond charged states may lead to precisely such non-perturbative effects.

Following the conjecture of Banks, Fischler, Shenker and Susskind [3], and using the result of Taylor [4], string configurations have been identified in 2d maximally supersymmetric Yang-Mills theory [5–7]. For another approach to string configurations in Yang-Mills theory, see [8]. What is remarkable about these models [5–7] is that weak string coupling is related to strong Yang-Mills coupling, most explicitly stated in [7]. Strong-coupling expansions typically have a finite radius of convergence [9], so it is of interest to identify how strong perturbation theory differs from Matrix string perturbation theory.

To address this issue, it is necessary to identify how the interactions of Matrix strings arise from the underlying Yang-Mills theory.

For concreteness, we focus on the identification of string configurations in the limit of vanishing string coupling proposed by Dijkgraaf, Verlinde and Verlinde [7]. An equivalent description of these configurations was given in [5,6]. Recall that, according to Taylor [4], Matrix theory [3] compactified on a circle is described by a 2d sYM theory on $R \times$ the dual circle. The action is just the dimensionally reduced action obtained from 10d sYM theory. The 8 components of the 10d gauge potential become scalar matter fields on the 2d worldsheet, corresponding to the the transverse space coordinates of the string in lightcone gauge. The action of interest is

$$S = \frac{1}{2\pi\alpha'} \int \text{Tr} \left( D_\mu X D^\mu X + \Theta T^{\mu\nu} D_\mu \Theta + \frac{1}{g_s^2} F_{\mu\nu}^2 \right) + \frac{1}{g_s} \left[ X', X' \right]^2 + \frac{1}{g_s} \Theta T' \left[ X, \Theta \right],$$

where $i = 1, \ldots, 8$, $\mu = 0, 1$, are worldsheet indices, and $\Theta$ is a Majorana-Weyl spinor in 10d.
worldsheet is taken to be cylindrical with \( \sigma \in [0, 2\pi] \).
This is to be considered for large \( N \), where \( U(N) \) is
the gauge group, following [3].

It was argued in [5–7] that at \( g_s = 0 \) the equations of
motion imply that the \( X \) matrices mutually commute,
with \([\Theta, X] = 0\), and hence may be diagonalized simultaneously.
The limit \( g_s \to 0 \) should correspond to an infrared fixed point, and hence to
a superconformal field theory. As one goes from \( \sigma = 0 \) to
\( \sigma = 2\pi \), the eigenvalues may have nontrivial
monodromy, in the sense that the eigenvalues only
label orbits of \( U(N) \) up to permutations, so if \( \lambda' \)
stands for the diagonalized \( X' \) matrices,
\[
\lambda(\sigma = 0) = P \lambda(\sigma = 2\pi) P^{-1},
\]
for some permutation matrix \( P \) in the defining rep-
resentation of \( S_N \). Refs. [5–7] suggested that string
configurations should correspond to cycles in this
permutation.

The interaction of strings in the model was ex-
plained in [7] as arising from the restoration of a
non-Abelian U(2) subgroup when two eigenvalues coincide. Surprisingly, the description of this interac-
tion given in [7], in terms of a twist field in the
superconformal field theory that describes the IR
fixed point at \( g_s = 0 \), makes no reference to the
Yang-Mills Lagrangian. Of course, if there is no contact with the Yang-Mills Lagrangian, the Matrix
approach would be a particularly obtuse way of
thinking about light-cone string perturbation theory
(which is not an enlightening approach to string
theory in the first instance). In particular, it is un-
clear why configurations of a few long strings should dominate the dynamics from the given description of
string interactions. The only non-perturbative content
in the model is the Yang-Mills action, and one would
like to derive the conjectured string interaction
from this action. This should also dynamically determine
which configurations of strings actually dominate the
dynamics.

We show in this note how the Yang-Mills gauge
field on the worldsheet determines the dominant
transitions between different string configurations in
this model. In doing so, we find that the eigenvalue
description is really only suited to \( g_s = 0 \). Away
from \( g_s = 0 \), the physics is much clearer in terms of
the full matrices. We believe this is the reason why
string perturbation theory diverges badly—it is sim-
ply that the description in terms of just string config-
urations, i.e. the eigenvalues in [5–7], is valid only at
\( g_s = 0! \)

We are interested in the sYM theory at strong
Yang-Mills coupling. Expanding \( S \) about an ultralo-
tical theory, we observe that, due to supersymmetric
cancellations, the first quantum corrections to the
action appear at order \( g_s^2 \) for configurations with
commuting \( X \) matrices. It is therefore meaningful to
consider the effects of classical terms of order \( g_s^0 \).
These terms imply that if the \( X \) matrices commute,
\( A \) is that connection such that the \( X \) matrices are
covariantly constant. Since the \( X \) matrices are sec-
tions of a twisted bundle, with monodromy \( P \), the
holonomy of the gauge connection must also be \( P \)
for minimizing the \( D_\mu XD^{*\mu}X \) term in the action.
While we expressed this in the continuum formula-
ation, it is even simpler to see this from the form of
the lattice covariant derivative, \( D_\mu X(x) \equiv U_{x,\mu}X(x) + e_\mu U_{x,\mu}^{-1} X(x) \).

Since the 2d sYM theory is strongly-coupled, an
appropriate starting point for calculations is the lat-
tice theory. Continuum weak-coupling engineering
dimensions of operators are not meaningful in such a
lattice theory, so we will not be able to derive the
dimension of the operator in the conformal theory
(\( g_s = 0 \)) that generates interactions. It is possible,
however, to ask qualitative questions about the na-
ture of interactions implied by the Yang-Mills action.

We deduce the approximate form of the dominant
transitions from the remaining non-vanishing term in
\( S \): The transition from a configuration described by a
monodromy \( P \) to one with monodromy \( P' \) is
weighted approximately (in Euclidean lattice gauge
theory terminology) by
\[
\exp \left( \frac{g_s^2}{2\pi\alpha'} \text{Tr}(P'P^{-1} + PP'^{-1} - 2) \right)
\]
in \( A^0 = 0 \) ‘gauge’. Thus the curvature of the gauge
field induced by a transition from one monodromy to
another is the determining factor in the relative
strengths of different interactions. When the mon-
odromy matrices are permutation matrices, it is easy
to see that the least possible curvature is induced by
a transition in which two cycles coalesce to form a
longer cycle, or vice versa. Comparing this to the
interactions described in [7], we see that the sYM
action provides a qualitative justification for the assumed form of string interactions.

This is, of course, a gross oversimplification, and should not be taken to mean that $F^2$ is the continuum operator that generates interactions in the strong-coupling conformal field theory. Indeed, the spectrum of scaling operators at $g_s = 0$ has little to do with the different terms in $S$. We have only deduced the qualitative form of the interactions of string configurations from the sYM action. The other terms that appear at the same order ($g^2$) are (1) quantum corrections to the ultralocal theory, and (2) quantum corrections from fluctuations of the terms that determined the holonomy of the connection. To determine the precise operator form of interactions at $g_s \neq 0$, one must include these quantum effects. In particular, the transition from one string configuration to another is affected by off-diagonal matrix elements of $X$, and by the fluctuations of the gauge field on the world-sheet. The dimension 3 string splitting operator [7] that appears exactly at $g_s = 0$ is a result of all these dynamical effects. For $g_s > 0$, one is dealing with the full matrix structure of the theory, and the stringy nature is no longer obvious.

For a more systematic exploration of the strong-coupling limit of the lattice gauge theory, it would be appropriate to rescale the fields in the action. Written as a dimensional reduction of 10d sYM theory, we would expect just to see $g^2$ in front of the entire action. By rescaling $X, \Theta$ one can write the action as $S$ above, or as

$$S' \equiv \frac{1}{2\pi\alpha'} \int \text{Tr} \left[ g_s D_\mu X D^\mu X + \sqrt{g_s} \Theta T^\mu D_\mu \Theta \right. $$

$$\left. + g_s^2 F_{\mu\nu} F^{\mu\nu} + \left[ X', X'' \right]^2 + \Theta T^\mu \left[ X, \Theta \right] \right],$$

a form suitable for an expansion in $g_s$ since all terms involving derivatives have positive powers of $g_s$. This is equivalent to a strong-coupling expansion in the Yang-Mills coupling.

The lack of a clear stringy interpretation for Matrix string interactions is a good thing, since if the stringy description were to be valid at $g_s \neq 0$, we would be left with firm evidence against Matrix theory, given the different perturbative behaviours of strong coupling Yang-Mills theory, describing Matrix strings, and garden-variety string theory. If one could surmount the problem of formulating supersymmetric lattice gauge theories [10], one would have a concrete tool to understand Matrix string theory, and see how its behaviour differs from string theory.

Conversations with Wati Taylor and Larus Thorlacius are gratefully acknowledged. This work was supported in part by NSF grant PHY96-00258.

References

Supergravity in 10 + 2 dimensions as consistent background for superstring

Hitoshi Nishino

Department of Physics, University of Maryland, College Park, 20742-4111, USA

Received 4 April 1997; revised 5 February 1998

Editor: M. Dine

Abstract

We present a consistent theory of $N = 1$ supergravity in twelve-dimensions with the signature (10,2). Even though the formulation uses two null vectors violating the manifest Lorentz covariance, all the superspace Bianchi identities are satisfied. After a simple dimensional reduction to ten-dimensions, this theory reproduces the $N = 1$ supergravity in ten-dimensions, supporting the consistency of the system. We also show that our supergravity can be the consistent backgrounds for heterotic or type-I superstring in Green-Schwarz formulation, by confirming the fermionic $\kappa$-invariance of the total action. This theory is supposed to be the purely $N = 1$ supergravity sector for the field theory limit of the recently predicted F-theory in twelve-dimensions. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

There has been some evidence [1] that type IIB and type I or heterotic string theories with no direct M-theory unification in 10 + 1 dimensions (11D) [2], may arise from a unifying theory called F-theory in 10 + 2 dimensions (12D). Like M-theory, F-theory is supposed to provide unified framework for understanding the vacuum structures of those string theories not unified by M-theory [3]. Motivated by this observation, we have established in our previous paper [4] the first superspace formulation for supersymmetric Yang-Mills theory in 12D with the signature (10,2). In this formulation we have presented the set of constraints consistent with the Bianchi identities (BIs) in superspace, together with the component results, performing also dimensional reductions into the conventional 10D and 4D supersymmetric Yang-Mills theories. More recently there has been super $(2 + 2)$-brane action formulated in flat superspace [5], and purely bosonic sector of F-theory has been also proposed [6]. However, the most important curved supergravity background in 12D consistently coupled to the above-mentioned string theories has been still lacking.

In this paper we take the first significant step toward the formulation of supergravity in 12D with two time coordinates. We will present in this paper the superspace formulation of supergravity in 12D with the manifest Lorentz covariance. The existence of such supergravity theory had been suggested for a long time in different contexts in the past because of small size of Majorana-Weyl spinors in such high dimensions as 12D [7], indicating the possibility of a
boson-fermion matching, or as possible $SO(10,2)$ covariant supergravity in the context of recent development of S-theory [8], or higher-dimensional theories with two time coordinates [9]. In this paper we will present such supergravity theory in an explicit but amazingly simple way, with the field representations for the algebra of $N = 1$ local supersymmetry in 12D. We will see how the system is avoiding possible inconsistency with broken Lorentz symmetry, while keeping some components in the extra dimensions non-vanishing, making the theory non-trivial. Our formulation is also similar to the globally supersymmetric Yang-Mills theory in 12D [4].

As important supporting evidence of the validity of our formulation, we will show how a simple supersymmetric Yang-Mills theory in 12D 4. As we introduce the Dirac matrices satisfying $(\sigma_{\mu}, \sigma_{\nu}) = 2 \eta_{\mu \nu} = 2 \text{diag}(\ldots, +, +, +, +, +, +, +, +, +, +, +)$, we get related important relations such as $\hat{n}\hat{\mu} + \hat{\mu}\hat{n} = 2I$, $\hat{n}^2 = \hat{\mu}^2 = 0$, etc. for the combinations $\hat{n} \equiv \sigma^a n_a$, $\hat{\mu} \equiv \sigma^a m_a$. It is then useful to define the projection operators

$$ P_\uparrow \equiv \frac{1}{2} \hat{n}\hat{n}, \quad P_\downarrow \equiv \frac{1}{2} \hat{\mu}\hat{\mu}, \quad (2.5a) $$

$$ P_\uparrow P_\downarrow = +P_\uparrow, \quad P_\downarrow P_\uparrow = +P_\downarrow, \quad (2.5b) $$

Depending on purposes, we sometimes switch the indices for the extra dimensions from $a,b,\cdots$ to $+,-$. A useful combination frequent in our formulation is

$$ P_\uparrow P_\downarrow = P_\uparrow - P_\downarrow = \sigma^{+-}. \quad (2.6) $$

It is also important to note the symmetry

$$ (\hat{n})_{a\beta} = -(\hat{n})_{\beta a}, \quad (\hat{\mu})_{a\beta} = -(\hat{\mu})_{\beta a}. \quad (2.7) $$

with the Majorana-Weyl spinor indices $\alpha, \beta, \cdots = 1, 32$ and $\dot{\alpha}, \dot{\beta}, \cdots = 1, 32$ in 12D. There are other important resulting identities, such as

---

2. Preliminaries with null vectors

We introduce two constant null vectors [1,4,5] in our $10 + 2$ dimensions with the signature $(-, +, +, \cdots, +, -)$, defined by

$$ (n^a) = \left(0,0, \cdots, 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), $$

$$ (n_a) = \left(0,0, \cdots, 0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), $$

$$ (m^a) = \left(0,0, \cdots, 0, \frac{1}{\sqrt{2}}, +\frac{1}{\sqrt{2}}\right), $$

$$ (m_a) = \left(0,0, \cdots, 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right). \quad (2.1) $$

with the local Lorentz indices $a,b,\cdots = (0),(1), \cdots, (9),(11),(12)$. These null vectors live only in the extra two dimensions for the coordinates $(11),(12)$ with the signature $(+, -)$.

Relevantly, it is convenient to define $\pm$ components associated with the extra dimensions for an arbitrary vector $V_a$:

$$ V_\pm \equiv \frac{1}{\sqrt{2}}(V_{(11)} \pm V_{(12)}) \quad (2.2) $$

Now it is clear that the factor $\sqrt{2}$ in (2.1) is to make the identifications such as

$$ n_+ = n^- = m_+ = m^- = 1, $$

$$ n_- = n^+ = m_+ = m^- = 0. \quad (2.3) $$

valid without extra $\sqrt{2}$ factor. Accordingly we have

$$ n^a n_a = m^a m_a = 0, $$

$$ m^a n_a = m^+ n_+ = m^- n^- = 1. \quad (2.4) $$

If we have only integral components in (2.1), the price we have to pay is the involvement of $2$ in (2.3).
The only difference of (3.1) from the usual BIs is the usage of the modified Lorentz generator $\tilde{M}$ satisfying
\[
(\tilde{M}_{ab})^c_d \equiv +\tilde{\delta}_{(a}^c \tilde{\delta}_{b)}^d
\times(\tilde{\delta}_a^c = 0, \text{otherwise } \tilde{\delta}_a^c = \delta_a^c),
\]

(3.3a)

\[
(\tilde{M}_{ab})_{(a}^c \equiv +\frac{1}{2} (\sigma_{ab} P_\gamma)_{a}^c \beta\gamma,
\]

(3.3b)

where the first one is similar to the usual Lorentz generator, while the second line has extra $P_\gamma$ and $P_\gamma$. We will clarify why this structure is needed to satisfy the BIs shortly. It is straightforward to confirm that the generators $\tilde{M}_{ij}$ with the 10D indices $i, j, \ldots = (0,1), \ldots, (9)$ form the SO(9,1) Lorentz algebra: $[\tilde{M}_{ij}, \tilde{M}_{kl}] = -\delta_{it} \{-1, 1\}_{jk}$, while $\tilde{M}_{_i}^j$ vanishes identically. Even though the commutators between $\tilde{M}_{++}$ and $\tilde{M}_{+i}$ does not obey the SO(10,2) Lorentz algebra, this poses no problem, as will be explained after Eq. (3.10). It is also straightforward to confirm that (3.3b) satisfies the Jacobi identity.

Our field content in 12D is the zwölfbein $e_m^a$, the gravitino $\psi_m^a$, the dilatino $\bar{\psi}_m$, antisymmetric tensor $B_{mn}$, and the dilaton $\Phi$. Especially, the gravitino (or the dilatino) has the negative (or positive) chirality: $\sigma_{ij} (\psi_m^a \bar{\psi}_n) = (\psi_m^a \bar{\psi}_n)$. The absence of the dotted gravitino $\bar{\psi}_m^a$ in our $N = 1$ system is related to the absence of the dotted indices $\tilde{a}, \tilde{b}, \ldots$ in (3.1) and (3.2).

Our results for superspace constrains are summarized as:

\[
T_{ab}^\gamma = (\sigma^{cd})_{\alpha_\beta \gamma} n_d + (\sigma^{cd})_{\alpha_\beta \gamma} n_d\gamma n_a m_b, \quad \gamma = 0,1, \ldots, 9, 11, 12, \text{essentially the same way as in [13].}
\]

To be more precise, there is an additional set of BIs (3.11) to be discussed shortly.

\# For curved bosonic coordinates we use $m, n, \ldots = 0, 1, \ldots, 9, 11, 12$, essentially the same way as in [13].

This chirality convention is the same as in [4]. We use, however, the bars for dotted spinorial superfields in this paper.
\[ T_{ab}^c = -G_{ab}^c, \quad (3.4g) \]
\[ R_{a\beta c\delta} = + \left( \sigma^{\epsilon \gamma} \right)_{a \beta} G_{c \delta}, \quad (3.4h) \]
\[ \nabla_a G_{bc \delta} = -\frac{1}{2} \left( \sigma^{\eta \xi} \nabla_a T_{\eta \xi} \right)_\delta = -\nabla_a T_{bc \delta}, \quad (3.4i) \]
\[ \nabla_a R_{ab \beta} = + \left( \sigma \partial_{a \beta} \right)_{\alpha} G_{ab \beta}, \quad (3.4j) \]

Here the spinorial inner products are the usual ones, e.g., \( \langle \bar{\chi} n T_{ab} \rangle = \bar{\chi}^\alpha (\bar{\eta})_\alpha n T_{ab}^\beta \). All other implicit components such as \( G_{a \beta} \) are zero, as usual. More importantly, due to the chiral nature of our superspace, all other chiral components, such as \( T_{ab}^c \) or \( T_{ab}^c \), are vanishing. Note, however, the existence of the component \( R_{A \beta}^\delta \) defined by \( R_{A \beta}^\delta = (1/2) R_{A \beta c} (\bar{\eta} \bar{\chi})_\delta \), needed for the computation of commutators such as \( \left[ \nabla_a, \nabla_b \right] \bar{\chi} = R_{A \beta}^\delta \bar{\chi}_\delta \). Note also that the components \( R_{A \beta}^\delta \) do not enter the BIs (3.1), because (3.1) is based on the identity \( \left[ \nabla_a, \nabla_b \right] + (2 \text{ perms.}) \equiv 0 \), with no derivative \( \nabla_a \) involved. This feature is also related to the fact that the dotted superfield \( \bar{\chi}_a \) is not defined like \( \nabla_a \phi \), but there is an operator \( \bar{n}_a \) multiplied in (3.4d), avoiding \( \nabla_a \).

In our system, there is further a set of extra constraints dictated by
\[ T_{a \beta}^c n_c = 0, \quad G_{a \beta} n^c = 0, \quad T_{a \beta}^c n_c = 0, \quad (3.5) \]
\[ R_{A \beta c} n^c = R_{A \beta c}^\delta n^\delta = 0, \quad (3.6) \]
\[ n \nabla_a \phi = 0, \quad n \nabla_a \bar{\chi}_a = 0, \quad (3.7) \]
\[ (\hat{m})_a \phi \bar{\chi}_a = 0, \quad T_{\alpha \beta} \gamma (\hat{m})_\gamma = 0, \quad (3.8) \]
\[ \phi_{a \beta} n_a = \phi_{a \beta} n_a = 0. \quad (3.9) \]

These are locally supersymmetric analog of the constraints \( F_{a \beta} n^0 = 0, \bar{\lambda} = 0, n \nabla_a \bar{\lambda} = 0 \) imposed in our previous paper [4]. Their role is to delete some (but not all) extra components that are non-physical after dimensional reductions into lower dimensions. By this prescription, we can get the right chiral field content for the resulting 10D supergravity, as will be seen shortly. It is also to be stressed that these extra constraints will not eliminate all the freedom of relevant superfields in the extra dimensions. For example, even though the component \( G_{a \beta} \) is eliminated by (3.5), the other extra component \( G_{a \beta} \) is still alive in our theory. This feature is crucial to have non-trivial system in 12D that is distinct from merely a rewriting of 10D system. This is similar to the supersymmetric Yang-Mills case [4].

Compared with our supersymmetric Yang-Mills theory in 12D [4], there is an extra term in (3.4a) which is needed to satisfy the \( d = 3/2 \) \((a b c d), (a b c d), (a b c d)\) BIs. Since this term is proportional to \( n^2 \), it has no effect in the commutator \( \left[ \nabla_a, \nabla_b \right] \propto \text{physical superfields} \), but it has new effect, only when \( T_{a \beta} \) is involved in the terms like \( T_{a \beta} \gamma T_{a \beta} \) in the \((a b c d)\) BI mentioned above. It is also important to note that the presence of this term will not affect our previous supersymmetric Yang-Mills system [4].

The structure (3.3a) is crucial for the covariant constancy of the null vectors:
\[ \nabla_a n^- = \partial_a n^- + \frac{1}{2} \phi_{a \beta} (\bar{\lambda} \gamma)_{a} n_{\alpha} = 0, \]
\[ \nabla_a m^+ = \partial_a m^+ + \frac{1}{2} \phi_{a \beta} (\bar{\lambda} \gamma)_{a} m_{\alpha} = 0. \quad (3.10) \]

Especially, the last terms vanish due to \( \bar{\delta}_{-} = 0 \) in (3.3). Note that we are not imposing conditions like \( \phi_{M} m_{\beta} = 0 \) directly, that would delete all the extra components in \( \phi_{M} \), reducing the theory to 10D supergravity. As has been mentioned, the irregular feature of the commutator \( \left[ \bar{\lambda}, \bar{\lambda} \right] \) poses no inconsistency, because when they are multiplied by the Lorentz connection, no extra component will be effectively left over. In fact, consider \( \phi_{a \beta} \bar{\lambda} \gamma \bar{\lambda} = \phi_{a \beta} \bar{\lambda} \gamma \bar{\lambda} \gamma + 2 \phi_{a \beta} \bar{\lambda} \gamma \bar{\lambda} \gamma + 2 \phi_{a \beta} \bar{\lambda} \gamma \bar{\lambda} \gamma \), where the last term vanishes due to \( \bar{\lambda} = 0 \), while the second and third terms also disappear thanks to (3.9). Additionally, due to the extra constraints (3.6), the commutators \( \left[ \nabla_a, \nabla_b \right] \propto \text{physical superfields} \) acting both on \( m \) and \( n \), vanish consistently. This system cleverly avoids inconsistency within the 10D sub-manifold, while maintaining the non-vanishing superfield components such as \( \phi_{M} \) in the extra directions.

In addition to the BIs in (3.1) and (3.2), there is another set of BIs called \( R \)-BIs of the type \((A B C D), \gamma \):\[ \nabla_{\gamma} R_{A B C D} \gamma = -T_{A B} R_{E F C D} \gamma \equiv 0. \quad (3.11) \]

Usually these BIs are automatically satisfied, once (3.1) holds [14]. In our system, however, this is non-trivial due to our modified Lorentz generators.
(3.3). Fortunately, we can easily confirm that the \((a_\beta y_d, e)\) BI at \(d = 3/2\) is satisfied, while for the \((a_\beta c, d, e)\) BI at \(d = 2\), we can show that the lemma used in Ref. [14] is still valid even with our modified Lorentz generators by explicit computation. The remaining \(R\)-BIs at \(d \geq 5/2\) will not give any non-trivial consistency checks. For example, we can confirm easily that the derivative of 3.8 vanishes:

\[
\left( a^{bc} \right)_{a \tilde{\beta} \, b \, \tilde{\gamma}} \left( \partial_\tilde{\gamma} \chi_{\tilde{\alpha}} \right) = 0 .
\]  

At \(d = 2\), (3.4j) will satisfy the \((abc, \delta)\) BI. Using this and taking a spinorial derivative of (3.13) in the combination \((\sigma_\delta)_{\tilde{\alpha} \tilde{\beta}} \left( \partial_\tilde{\gamma} \chi_{\tilde{\alpha}} \right)\)

\[
\left[ R_{\alpha \beta} \delta_{\gamma} \right] + 4 \left( \partial_\tilde{\gamma} \chi_{\tilde{\alpha}} \right) \left[ R_{\alpha \beta} \delta_{\gamma} \right] = 0 .
\]  

Finally, the important relationship

\[
R_{\alpha \beta} = - \nabla_a G_{\alpha \beta} \]  

comes out of the \((abc, d)\) BI via the contraction \(X_{\alpha \beta c} \). All of these are just parallel to the 10D case [10], and are also analogous to the corresponding ones for supersymmetric Yang-Mills multiplet in [4].

Some remarks are in order for these results. There seems to be a priori no systematic method to fix the constraint system we obtained in this paper, due to the lack of Lorentz covariance inherent in the system. Our important guiding principle that led us to these expressions is to reproduce the well-known \(N = 1\) supergravity in 10D after simple dimensional reduction to be explained shortly. Before reaching our results (3.4) and (3.5), we have tried many different options, such as assigning the same chirality in 12D for \(\psi_m \) and \(\chi_a \), modifying the \(R\) or \(G\)-BIs by Chern-Simons forms, or requiring no \(G\)-BI at all, assuming that the dilaton and the dilatino in 10D would come out of the extra components in the zwölfein and the gravitino, all in vain. For example, the idea of getting the dilatino out of the 12D gravitino failed, because the \(\chi\)-transformation rule did not arise right out of the gravitino in the dimensional reduction process we deal with next. All of these trials seem to indicate the dimensional reduction is a key guiding principle to fix the right constraint set in 12D. It is a kind of “oxidation” procedure from 10D to 12D that led us to our results. In particular, fixing the right form for \(T_{\alpha \beta} \) seems to be the crucial step. Relevantly, we found that the modified generator (3.3) for the dotted indices are crucial, in order to get the right gravitino/dilatino superfield Eq. (3.13) at \(d = 3/2\) that reproduces 10D superfield Eq. (4.8) to be discussed.
Compared with the supersymmetric Yang-Mills case in [4], we noticed that no auxiliary fields, such as the $\chi$-field (in the notation of Ref. [4]) are needed for the satisfaction of BI's. We do not know yet the reason why we need no such auxiliary fields for supergravity.

Before ending this section, we give the explicit component field supersymmetry transformation rule obtained from our constraints (3.4) using the method in Ref. [13]:

$$\delta_0 e_m^a = (e\sigma^{ab}\psi_m)n_b + (eP_+^a\psi_m)n^a,$$

$$\delta_0 \Phi = - (e\check{\eta}\check{\chi}),$$

$$\delta_0 \psi_m^a = D_m e^a + (P_+ e)^a(\check{\chi}\check{\eta}\psi_m)$$

$$+ (P_+\check{\psi}_m)^a(e\check{\eta}\check{\chi})$$

$$- (P_+\sigma\check{\psi}_m)^a(e\check{\eta}\check{\chi})n_b,$$

$$\delta_0 B_{mn} = + (e\sigma^m\check{\psi}_n)n_b - (eP_+^a\psi_m)n^a,$$

$$\delta_0 \check{\chi}_a = + \frac{i}{2} (P_+ \sigma^{mn}e)_a G_{mn},$$

$$+ \frac{i}{2} (P_+ \sigma^{mn}e)_a \gamma^a \Phi - \check{\chi}_a (e\check{\eta}\check{\chi}). \tag{3.16}$$

The Lorentz connection $\phi^{ab}_m$ involved in $D_m e^a \equiv \partial_m e^a = (1/4)\phi^{ab}_m (\sigma^a P_+ e)^b$ has the torsion $T_{ab}^c$ as well as the $\psi$-torsion with our $T_{a\beta}^\gamma$ in (3.4a) [13]. Compared with the supersymmetric Yang-Mills case [4], the transformation of $e_m^a$ now has the null vector, while the leading term in $\delta_0 e_m^a$ does not.

4. Dimensional reduction into 10D

Next important confirmation is to show that our system reproduces already known results in lower dimensions such as 10D. In this paper we perform a simple dimensional reduction from 12D to 10D, paying special attention to the null vectors we introduced. Other than the role played by the null vectors, our prescription of dimensional reduction in superspace is similar to that in Ref. [15].

It has been known that there are innumerable many mutually equivalent sets of constraints for $N = 1$ supergravity in 10D due to the freedom of super Weyl-rescaling. For simplicity of computation, we use what is called ‘‘beta-function-favored constraints’’ (BFFC) [10], which is the simplest set of constraints among possible constraint sets, arranged by appropriate super Weyl-rescalings. By this choice, our constraint system is drastically simplified in 10D, that saves considerable effort in our dimensional reduction.

To distinguish the 10D quantities from the original 12D ones, we introduce the hat symbols on the fields and indices in 12D, only within this section. This procedure is a supergravity analog of the similar one we have performed in Ref. [4]. We first setup the dimensional reduction rule for the $\sigma$-matrices as in [4], as

$$\hat{\sigma}_a = \begin{cases} \hat{\sigma}_a = \sigma_9 \otimes \tau_3, \\
\hat{\sigma}_{(11)} = I \otimes \tau_1, \\
\hat{\sigma}_{(12)} = -I \otimes i\tau_2, \end{cases} \tag{4.1}$$

where $\tau_i (i = 1, 2, 3)$ are the usual Pauli matrices, and all the non-hatted quantities and indices are for 10D. Accordingly, the dimensional reduction for the charge conjugation matrix $\hat{C}$ and $\hat{\sigma}_{13}$ in 12D are [4]

$$\hat{C} = C \otimes \tau_1, \quad \hat{\sigma}_{13} = \sigma_{11} \otimes \tau_3, \tag{4.2}$$

with the charge conjugation matrix $C$ in 10D. Subsequently, $\check{\eta}$ and $\eta$ satisfy the relationships such as

$$\left(\hat{\eta}^\dagger\right)_{\hat{a} \hat{b}} = \left(\check{\eta}^\dagger\right)_{\check{a} \check{b}} = \sqrt{2} I \otimes \begin{pmatrix} 0 & 1 \\
0 & 0 \end{pmatrix},$$

$$\left(\hat{\eta}^\dagger\right)_{\hat{a} \hat{b}} = \left(\check{\eta}^\dagger\right)_{\check{a} \check{b}} = \sqrt{2} I \otimes \begin{pmatrix} 0 & 0 \\
1 & 0 \end{pmatrix},$$

$$\hat{P}_+ = I \otimes \begin{pmatrix} 1 & 0 \\
0 & 0 \end{pmatrix}, \quad \hat{P}_- = I \otimes \begin{pmatrix} 0 & 0 \\
0 & 1 \end{pmatrix}. \tag{4.3}$$

In this representation, the operations of $\hat{P}_+$ and $\hat{P}_-$ are transparent. We next require that the $\hat{\chi}$ and the gravitino field strength $\hat{T}_{ab}$ have the components

$$\left(\hat{\chi}_{ab}\right) = \begin{pmatrix} \hat{\chi}_{a \check{a}} \\
\hat{\chi}_{a \check{a}} \end{pmatrix} = \begin{pmatrix} 0 \\
\chi_a \end{pmatrix},$$

$$\left(\hat{T}_{ab}^\gamma\right) = \begin{pmatrix} \hat{T}_{ab}^\gamma \chi_{a \check{a}} \\
\hat{T}_{ab}^\gamma \chi_{a \check{a}} \end{pmatrix} = \begin{pmatrix} T_{ab}^\gamma 0 \end{pmatrix}. \tag{4.4}$$
where $X_{\mu}$ and $T_{a\beta}^\gamma$ are to be the resulting 10D superfields. From now on, we use the index $\uparrow$ (or $\downarrow$) for the first (or second) component of a 12D spinor decomposed into the Pauli matrix space in (4.1). These ansätze are consistent with our extra constraints (3.5)–(3.9), as well as their 12D chiralities, via (4.2), which had been fixed in such a way that the resulting 10D supergravity theory has the right chiralities.

We can now easily show that all the constraints (3.4) and the superfield Eqs. (3.13)–(3.15) are reduced into 10D following the simple dimensional reduction in superspace [15], respecting also $\tilde{\Phi}_s = \tilde{\Phi}_c = 0$, $\tilde{R}_{+B}^{} = \tilde{R}_{A+}^{} = 0$, $G_{+A} = 0$, as usual.

For example, the dimensional reduction for the $\tilde{c} = c$, $\tilde{\alpha} = \alpha \uparrow$, $\tilde{\beta} = \beta \uparrow$ -component of $\tilde{T}_{\alpha\beta}^\gamma$ becomes

$$
\tilde{T}_{\alpha\beta}^\gamma \rightarrow T_{\alpha\beta}^\gamma
$$

in agreement with the 10D result [10], up to a non-essential rescaling factor $\sqrt{2}$. Similarly, if we look into the $\tilde{\alpha} = \alpha \uparrow$, $\tilde{\beta} = \beta \uparrow$, $\tilde{\gamma} = \gamma \uparrow$ -component of $\tilde{T}_{\alpha\beta}^\gamma$:

$$
\tilde{T}_{\alpha\beta}^\gamma \rightarrow T_{\alpha\beta}^\gamma
$$

again in agreement with the 10D result [10]. Parallel procedures work for other constraints in (3.4), reproducing BFFC [10]. We can also make it sure that all the superfield equations in 10D are reproduced by similar procedures from (3.13) and (3.14). In fact, by looking into the $d[\tilde{b}\tilde{c}] = d[b + c]$-component of (3.14), we get the 10D gravitational superfield equation in BFFC [10]:

$$
R_{ab} + 4\nabla_a \nabla_b \Phi - 4\sqrt{2} (\tilde{T}_{ab}^c \chi) = 0 ,
$$

while all other components such as $d[\tilde{b}\tilde{c}] = d[b - c]$ are trivially satisfied. Here we performed the usual identification $\tilde{R}_{ab} = R_{ab}^c \Phi = \Phi$, etc. The $G$-superfield equation is also contained herein by the relationship $R_{(ab)} = -\nabla_a G_{bc}^{}$ [10]. Similarly from the $\tilde{a} = a$, $\tilde{\alpha} = \alpha \uparrow$ -component of (3.13) we get the gravitino/dilatino superfield equation

$$
\sigma^c T_{ab}^c + \nabla_a \chi = 0 ,
$$

in agreement with [10].

5. Green-Schwarz superstring $\sigma$-model action

We have thus far established the superspace formulation of $N = 1$ supergravity in 12D. However, we still need to see if such a system can be consistent backgrounds for heterotic or type-I superstring theory. Here we give the action for Green-Schwarz superstring $\sigma$-model, and confirm its $\kappa$-invariance. Due to our extra coordinates, we need a peculiar constraint lagrangian, and we show how such term is consistent with the $\kappa$-symmetry. This confirmation provides supporting evidence for our system to be regarded as the 12D origin of heterotic or type-I superstring theories.

Our total action $I$ is composed of the $\sigma$-model action $I_{\sigma}$, the Wess-Zumino-Novikov-Witten term $I_{B}^{}$, and the constraint action $I_{\Lambda}^{}$:

$$
I \equiv I_{\sigma} + I_{B} + I_{\Lambda} ,
$$

$$
I_{\sigma} = \int d^3 \sigma \left[ V^{-1} \eta_{ab} \Pi_+^a \Pi_-^b \right] ,
$$

$$
I_{B} = \int d^3 \sigma \left[ V^{-1} \Pi_+^a \Pi_-^a B A \right] ,
$$

$$
I_{\Lambda} = \int d^3 \sigma \left[ V^{-1} \Lambda_{+}^a \Pi_-^a n_{a} (\Pi_-^b m_{b}) + V^{-1} \Lambda_{+}^a \left( (\Pi_-^a n_{a})^2 + (\Pi_-^a m_{a})^2 \right) \right] .
$$
where we use the zweibein \( V_+ \), \( V_- \) for the 2D world-sheet with the coordinates \( \sigma^i \), and \( V = \det(\text{something}) \) with the curved indices \( i, j, \ldots \) and the flat light-cone indices \( (i), (j), \ldots = +, - \), while \( \Pi^{ij} \equiv \frac{1}{2} \left( \delta Z^{ij} \right) E_{ij} \) with the vielbein \( E_{ij} \) for the 12D superspace with its supercoordinates \( Z^{ij} \). The action \( I_A \) has non-propagating Lagrange multipliers \( \lambda_{++} \) and \( \tilde{\lambda}_{++} \), deliberately chosen such that their field equations get rid of the unwanted contribution to the conformal anomaly, or to the \( \kappa \)-variation of our total action, as will be seen next.

Our action has two fermionic symmetries with the parameters \( \kappa \), the former of which is analogous to the 10D case [16,17]:

\[
\delta V_+ = \kappa (\sigma^i) \partial \Pi^{ij} n_i V_-, \quad \delta V_- = \kappa (\sigma^i) \partial \Pi^{ij} n_i V_+.
\]

By adding (5.7) to (5.6), it is clear that the variation of our total action vanishes:

\[
\delta \kappa I_A = \delta \kappa (I_2 + I_3 + I_4) = 0.
\]

In a similar fashion, using \( \delta \eta I_A = 0 \), etc., we can show the \( \eta \)-invariance

\[
\delta \eta I = 0,
\]

whose meaning will be explained shortly.

The significance of \( I_A \) is seen as follows. Without \( I_A \), we have to put a constraint \( \Pi^{-nn} = 0 \) or \( \Pi^{-mn} = 0 \) by hand for the \( \kappa \)-invariance. However, since this equation has a first derivative on the 2D world-sheet, such an equation is no longer regarded as a constraint, but is a "field equation" that should not be imposed upon the invariant check. We also mention that \( I_A \) or the extra bosonic coordinates in 12D will not contribute to the conformal anomaly in 2D world-sheet. Consider the field equations of \( \Lambda_{++} \) and \( \Lambda_{+++} \):

\[
(\Pi^{-nn}) (\Pi^{-nn} m_n) = 0,
\]

\[
(\Pi^{-nn})^2 + (\Pi^{-mn})^2 = 0,
\]

which are equivalent to

\[
\Pi^{-nn} m_n = 0, \quad \Pi^{-mn} m_n = 0.
\]

Eq. (5.11) physically implies that the extra coordinates \( X^\pm \) are independent of the world-sheet coordinate \( \sigma \). Once \( \Pi^{-nn} \) vanishes, \( \Pi^{-nn} \) in \( I_2 \), \( I_3 \) or \( I_4 \) will not contribute to the energy-momentum, and therefore not to the conformal anomaly. The contribution of \( I_2 \) to the \( X^\pm \)-field equation also vanishes, due to the factor \( \Pi^{-nn} \) always involved, while \( \Lambda_{++} \) and \( \Lambda_{+++} \) themselves disappear from field equations, and therefore will not enter the classical string spectrum. This feature is exactly the same as the constraint actions in Ref. [18]. Additionally we can see that all the extra components in the \( X^\pm \)-field equations are now satisfied under (5.11).

Finally the \( \eta \)-invariance with (5.5c) implies the redundancy of half of the total 32 components of the
superspace coordinates $\theta^\mu$ in 12D, like a gauge
symmetry. After all, thanks to the $\kappa$- and $\eta$-fermionic
invariances, the original 32 degrees of freedom of
$\theta^\mu$ are reduced like $32 \to 16 \to 8$, in accordance
with the 10D Green-Schwarz formulation [16].

6. Concluding remarks

In this paper we have presented supergravity the-
ory in 12D for the first time. We have shown how
the superspace BIs are satisfied in the presence of
null vectors. We have also confirmed that the BIs
yield the superfield equations in 12D. We next per-
formed the simple dimensional reduction into the
usual 10D, reproducing the BFFC system [10] for
10D, $N = 1$ supergravity, as good supporting evi-
dence for validity of our system.

As another crucial consistency check, we have con-
structed a non-trivial Green-Schwarz superstring
$\sigma$-model action that can couple to our 12D super-
gravity background. We found that the introduction
of the constraint action $I_\kappa$ guarantees the $\kappa$-invari-
ance of the total action. Our action $I_\kappa$ resembles the
constraint action discussed in [18] bilinear in con-
strained fields, and the multipliers disappear from
field equations. This mechanism also helps to main-
tain the cancellation of conformal anomalies on the
string world-sheet compared with 10D superstring.
We have also found the importance of the $\eta$-
fermionic invariance, that together with the $\kappa$-invari-
ance reduces the original 32 degrees of freedom of
$\theta^\mu$ into 8, in accordance with the 10D superstring
theories.

The field content of our system does not produce
the 11D supergravity, as seen from the absence of a
forth-rank field strength. Even though we skipped in
this paper, we can easily couple the supersymmetric
Yang-Mills multiplet [4] to our supergravity within
12D [19], by modifying the superfield strength $G_{\alpha\beta\gamma}$
by Chern-Simons forms [20]. Another direction is to
enlarge the number of supersymmetries to $N = 2$
[21] which is closer to the F-theory [1].

We stress that our theory is not merely a re-writ-
ing of the 10D, $N = 1$ supergravity, because extra
components like $G_{ab+}$ or $T_{a+}$ are non-vanishing in
12D. This is analogous to the supersymmetric
Yang-Mills case in 12D [4] with the gauge and
gaugino depending on the coordinate $X^+$, as has
been clarified in Section 3. Even though these extra
components are not physical, they play a role of
keeping the system distinct from merely a re-writing
of the 10D supergravity. They are like the third and
fourth components of a gauge field $A_\mu$ in 4D in the
Lorentz gauge, because even though they are merely
gauge degrees of freedom and disappear in the
light-cone, they play a formal role to make all the
four components of $A_\mu$ manifest. As different gauge
choices have different significances distinct from the
light-cone formulation, we emphasize that our 12D
formulation has its proper significance.

The formulation of supergravity theory in 12D
has now been made possible by the introduction of
null vectors. Our investigation was motivated by the
recent F-theory [1], (2 + 2)-brane [5], or S-theory [8]
that suggest supergravity theories made possible by
null vectors. The recent work for the globally super-
symmetric Yang-Mills theory [4] also indicated the
existence of supergravity with the superfield strength
$G_{\alpha\beta\gamma}$. Our explicit result provides new concept that
local supersymmetry can be formulated consistently
with the constant null vectors. Our result provides
strong motivation to explore other supergravity theo-
ries in 12D, in particular, formulating $N = 2$ super-
gravity in 12D for the field theory limit of F-theory
[1] as the underlying theory for type IIB superstring.
Armed with the working $N = 1$ supergravity system,
the construction of the 12D $N = 2$ supergravity must
be straightforward. Our result also suggests other
supersymmetric theories in dimensions $\geq 12D$, like
the recently-proposed supersymmetric Yang-Mills
theory in 14D [22].

Since we have established a supergravity theory
in curved 12D, we can explore other possible compac-
tifications directly from 12D into dimensions $\leq
10D$ [6]. Now with the gravitational field in curved
12D, it is easier to consider more sophisticated compac-
tifications. It is interesting to see if any new
feature arises from the extra dimensions in 12D that
did not show up in the compactifications of 10D
superstring theories.

The introduction of null vectors into superspace
itself is not an entirely new concept. In Ref. [10], we
have introduced similar null vectors for constraints
in superspace for $\beta$-function in the Green-Schwarz
$\sigma$-model. It seems that the necessity of these null
vectors in the Green-Schwarz formulation is inherent in superstring theories, which were originally formulated in the light-cone gauge. If we try to maintain the ‘‘covariance’’ as formally as possible in the Green-Schwarz formulation, these null vectors are to be involved, and our 12D theory is such an example.

The compatibility of our superspace backgrounds with Green-Schwarz superstring σ-model is also consistent with the prediction that heterotic or type I superstring may come from F-theory in 12D [1]. Despite of the lack of an invariant lagrangian in 12D, our Green-Schwarz formulation provides the target space effective action 23, like other similar superstring theories such as type IIB theory with no lagrangians. It is also interesting to see if the Green-Schwarz σ-model β-functions for our 12D theory provide our superfield equations [10]. Another direction is to explore a (2 + 2)-brane action [5] that may well be a more natural p-brane action for our 12D superspace. Studies for these directions are also under way.

Acknowledgements

We are grateful to S.J. Gates, Jr., E. Sezgin, W. Siegel, and C. Vafa for important discussions at many stages of our work presented in this paper.

References

New $N = 1$ supersymmetric 3-dimensional superstring vacua from U-manifolds

Gottfried Curio $^a$, Dieter Lüst $^b,1$

$^a$ School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540, USA
$^b$ Humboldt-Universität, Institut für Physik, Invalidenstraße 110, 10115 Berlin, Germany

Received 9 March 1998
Editor: L. Alvarez-Gaume

Abstract

Making use of non-perturbative U-duality symmetries of type II strings we construct new ‘superstring’ vacua in three dimensions with $N = 1$ supersymmetry. This has an interpretation as compactifying formally from 13 dimensions (S-theory) on Calabi-Yau 5-folds possessing a $T^3 \times T^2$ fibration. We describe some part of the massless multiplets, given by the Hodge spectrum, and point to a corresponding 5-brane configuration. © 1998 Published by Elsevier Science B.V. All rights reserved.

1. S-theory

The F-theory construction [1,2] can be generalized by considering the type IIB string compactified (on a torus) to lower ($d'$) dimensions [3]. If one allows the scalar fields (U-fields) to vary over some part of the $d'$-dimensional space and allows them to jump, consistent with the U-dualities, this data will translate to a (complex) n-dimensional manifold $K^*$ whose (real) $b$-dimensional base $B^b$ is the visible space from $d'$ to $d' - b$ dimensions and the (real) $2n - b$ dimensional fibre being the geometrization of the U-duality. Hence this construction leads to type IIB string vacua with $d = d' - b$ flat, uncompactified dimensions. As an example consider S-theory [3,4] with $d' = 8$. In this case the scalar moduli space is given by the coset

$$SL(3, Z) \times SL(2, Z) \times SO(3) \times SO(2),$$

(1)

and the U-duality group is $SL(3, Z) \times SL(2, Z)$. In S-theory the seven scalar fields, which parametrize the above coset, are allowed to vary over the 5-dimensional base $B^5$. The U-duality group arises as a combination of two contributions: on the one hand one has the $SL(2, Z)$ which exists already in 10 dimensions (the ‘S-duality’ of the type IIB string), and is used there to append the F-theory elliptic torus, leading to a theory living formally in 12 dimensions. This $SL(2, Z)$ is united with the $SL(2, Z) \times SL(2, Z)$, which arises on the other hand after compactification of the type IIB theory to 8 dimensions on a $T^2$; to be precise, $SL(2, Z)$ combines with the

---

$^1$ E-mail: curio@ias.edu., luest@qft1.physik.hu-berlin.de. The first author is partially supported by NSF grant DMS 9627351.
Sl(2, Z), to the Sl(3, Z), whereas the Sl(2, Z)
remains giving the Sl(3, Z) factor of Sl(3, Z)× Sl(2, Z). Note that af-
er compactification on the $T^2$ to 8 dimensions the theory becomes equivalent to type IIA and thereby to $M$
theory on $T^3$, which gives a further view on the U-duality group Sl(3, Z)× Sl(2, Z). So the (possibly reducible)
Calabi-Yau 5-fold $K^5$ must be a $T^3× T^2$ fibration over $B^5$:

$$K^5 \to \tau^3× \tau^2 B^5.$$  (2)

Since one has appended in 8 dimensions a 5-dimensional torus $T^3× T^2$, S-theory can be regarded as a
13-dimensional theory.

Next we discuss what kind of 5-folds $K^5$ with $T^3× T^2$ fibration can be constructed. First consider splitting
(reducible) 5-folds which lead to $N=2$ supersymmetry in 3 dimensions. One possible choice [3] is a product
space where one factor is a Calabi-Yau 3-fold $CY^3$ with $T^3$ fibration and the other factor is an elliptic $K^3$, i.e.

$$K^5 = CY^3× K^3, \quad CY^3 \to \tau^3 B^3, \quad K^3 \to \tau^3 S^2.$$  (3)

Therefore the total 5-dimensional base $B^5$ is given by

$$B^5 = B^3 × S^2.$$  (4)

We will always assume that $B^3 = S^3$ (cf. also [4]; this is connected with mirror symmetry on Calabi-Yau 3-folds
[5], [6]). Another class of vacua with $N=2$ supersymmetry is given by

$$K^5 = CY^4× T^2, \quad CY^4 \to \tau^3 B^5.$$  (5)

The $CY^4$ is assumed to be $CY^3$ fibered over $P^1$. Therefore $B^5$ is a $S^3$ fibration over $S^2$.

In this paper we will see (Section 3) that Calabi-Yau 5-folds exist which are $T^3× T^2$ fibrations but not
product spaces. This leads to $N=1$ supersymmetry in 3 dimensions. Before, this was, in string compactifica-
tion, possible to be reached only using the somewhat difficult to handle spaces of exceptional holonomy
(heterotic string on $G_2$-manifold, M-theory on Spin(7)-manifold). By contrast it is realised here still in the
framework of the well-suited Calabi-Yau spaces.

The new three-dimensional superstring vacua described in this paper might also lead to a geometric
understanding of (non-perturbative) effects in 3D, $N=1$ supersymmetric field theory [7], possibly very much
like the realisation of instantons, contributing to the superpotential, via internal geometrical cycles in the context
of $M$-theory resp. $F$-theory on a Calabi-Yau four-fold (leading to $N=2$ in 3D resp. $N=1$ in 4D).

Finally in view of Witten’s scenario [8] relating $N=1$ supersymmetric theories in three dimensions to
non-supersymmetric theories in four dimensions with vanishing cosmological constant one can try to relate the
described 3D vacua to 4D vacua of $N=0$ like it was tried [1] for $M$-theory on Spin(7) manifolds.

2. Some dualities

Let us first recall [3] the duality symmetries between S-theory on the one hand and F-theory and the heterotic
string on the other hand. One derives first that S-theory on $K^5× S^1$ is dual to F-theory on $K^5$:

$$d=2: \quad S|_{K^5× S^1} \leftrightarrow F|_{K^5}.$$  (6)

Here in $F|_{K^5}$ - i.e. as soon as one has leaved S-theory, which is in a sense a type IIB theory with additional
structure, and has reached F-theory, which is a type IIB theory with a different additional structure - the $T^2$
fibre of $K^5$ corresponds to the elliptic fibre used in F-theory to codify the type IIB complex coupling constant.
On the other hand, the volume of the $T^3$ fibre of $K^5$ in F-theory corresponds to the inverse radius of $S^1$. One can go further down in dimensions and arrives at the following chain of dualities

$$d = 1:\ S_{K^5 \times T^2} \leftrightarrow F_{K^5 \times S^1} \leftrightarrow M_{K^5},$$

(7)

where the inverse radius of the extra circle is related to the volume of the $T^2$ fibre in M-theory.

Before we discuss S-theory on $CY^5$ in greater detail let us consider the reducible cases of S-theory on $K^5 = CY^3 \times K3$ with an elliptically fibered $K3$ resp. on $K^5 = CY^4 \times T^2$ with a $CY^3$ fibered $CY^4$. In case of S-theory on $K^5 = CY^3 \times K3$ the above chain of dualities can be extended [3]. As, upon compactification on $S^1$, this is dual to F-theory on $CY^3 \times K3$ and F-theory on the elliptic $K3$ is dual to the heterotic string on $T^5$, this means that in 2 dimensions we have a duality with a heterotic string on $CY^3 \times T^2$:

$$d = 2:\ S_{CY^3 \times K3 \times S^1} \leftrightarrow H_{CY^3 \times T^2},$$

(8)

The other reducible case of S-theory on $K^5 = CY^4 \times T^2$ with a $CY^3$ fibered $CY^4$ is also interesting to consider and not already covered (via duality) by some other theory known before, especially it is not dual to $M$-theory on $CY^4$. For this note that after compactification on $S^1$ the $T^2$ factor is now the F-theory elliptic fibre, i.e. this is simply type IIB on $CY^4$ of $(0,4)$ spacetime supersymmetry in 2 dimensions. By contrast $M$-theory on $CY^4$ is the lifting to 3 dimensions of the non-chiral theory of $(2,2)$ space-time supersymmetry in 2 dimensions given by type IIA on $CY^4$; in other words the 13-dimensional S-theory on $CY^4 \times T^2$ can be viewed as the lift of type IIB on a $CY^4$ from 2 to 3 dimensions just as the 11-dimensional $M$-theory does the corresponding thing for type IIA.

Let us come now to the 3-dimensional theories with $N=1$ supersymmetry like S-theory on $CY^5$ or $M$-theory on a Spin(7) manifold (or the heterotic string on a $G_2$ manifold). Let us consider, in view of the observation just made for the reducible $CY^4 \times T^2$ case, again first the situation in 2 dimensions. There one has again the non-chiral theory given by type IIA on a Spin(7) manifold and the chiral one given by F-theory on $CY^5$. Then $M$ theory on Spin(7) lifts the first, non-chiral, theory to 3 dimensions, whereas S-theory on $CY^5$ lifts the chiral theory to 3 dimensions.

3. Some Calabi-Yau 5-folds

For the compactification of S-theory we will now construct (complex) Calabi-Yau 5-folds which have a $T^3 \times T^2$ fibration. The $T^3$ part will be always get from a $CY^3$ (which is assumed to have a mirror, and so a $T^3$ fibration over a real 3-dimensional base $B^3 = S^3$, cf. [5]). We concentrate on spaces which are true (irreducible, non-splitting) $CY^3$. We will discuss below examples of the form $X^4 \times_P dP_g$, where $X^4$ is a (non Calabi-Yau) four-fold which has a $CY^3$ fibration over $P^1$ (whose fibration structure is, which is also part of the ‘input data’ structure and not given by the $CY^3$ alone, will be described below) and

$$dP_g = \begin{bmatrix} p^2 & 3 \\ P^1 & 1 \end{bmatrix}$$

is a surface having an elliptic fibration over $P^1$. Clearly, the very idea of this fibre product is that the splitting of the fibration does not imply the splitting of the total space. The real 5-dimensional base $B^5$ for the $T^3 \times T^2$ fibration is a $S^3$ fibration over $S^2$, which will be described in more detail in the next section.

---

Footnote 1: For example one can have a description derived from a representation $CY^3 \cong (CY^4 \times T^2)/Z_3$ which is to be understood in the same sense as the construction of the $CY^{19,19} \cong dP_9 \times dP_9$ from the quotient $(K3 \times T^2)/Z_3$ (cf. appendix and [11]); similarly one could study a $CY^3 \cong (CY^3 \times K3)/Z_3$ version, this time with the $dP_g$-fibrefactor appearing by ‘reduction’ from $K3$.

Footnote 2: Cf. the appendix and [9–14]
For example one can build out of the Calabi-Yau 3-fold $CY^{19,19}$, by fibering it over a further $P^1$, the 4-fold

$$X^4 = \begin{bmatrix} p^2 & 3 & 0 \\ p^1 & 1 & 1 \\ p^2 & 0 & 3 \\ p^1 & 0 & 1 \end{bmatrix} = dP_g \times_{P^1} K$$

with

$$K = \begin{bmatrix} p^1 & 1 \\ p^2 & 3 \\ p^1 & 1 \end{bmatrix}$$

and finally the Calabi-Yau 5-fold

$$CY^5 = \begin{bmatrix} p^2_a & 3 & 0 & 0 \\ p^1_b & 1 & 1 & 0 \\ p^2_c & 0 & 3 & 0 \\ p^1_d & 0 & 1 & 1 \\ p^2_e & 0 & 0 & 3 \end{bmatrix} = dP_g \times_{P^1} K \times_{P^1} dP_g,$$

We will analyse that example 4 further below.

Let us analyse now the cohomology of a general 5 (i.e. at first not necessarily of the fibre-product form) Calabi-Yau 5-fold. Its Hodge diamond looks like

```
  1
 0 0 h^{11} h^{21} h^{31} 0
 0 0 0 h^{12} h^{22} h^{32} 0
 1 h^{41} h^{42} h^{43} h^{44} 1
```

Below we will compute the Hodge numbers which have the most immediate interpretation, $h^{11}$ as Kähler parameters and $h^{11}$ as complex deformations, from our input data in the class of $X^4 \times_{P^1} dP_g$ spaces. Furthermore from the $CY^3 \times T^2$ fibration over $P^1$ of these spaces one gets for the Euler number $e$ of the 5-fold as $e = 12 \cdot e(CY^3)$ as $e(dP_g) = 12$. So we will ‘know’ 3 of the 6 unknown numbers. In the Calabi-Yau 4-fold case one gets one further information from a relation derived in [15]; this is enough for the 4 unknowns in the 4-fold case, in our case 2 unknowns remain. Let us see in detail how this happens.

\footnote{Note that one has actually an $T^2 \times T^2 \times T^2$ fibration over $P^1 \times P^1$; in the quotient description here the $CY^4 = dP_g \times_{P^1} \mathcal{X}$ (cf. appendix) with $E_8$ superpotential of [10,11] occurs.}

\footnote{Assumed to be non-splitting; the Hodge diamonds for the reducible cases $CY^4 \times T^2$ and $CY^3 \times K3$ are of course trivially computed.}
The index of the $(1,0)$-forms-valued $\delta$ operator, $\text{ind} \, \delta = \sum_{q=0}^{5} (-1)^q h_{q+1}$, is according to the index theorem given by

$$\text{ind} \, \delta = \int_{X} Td(X) \, ch(T^*X),$$

where for the Calabi-Yau 5-fold $X$ one has

$$Td(X_{CY}) = 1 + \frac{e_2}{12} + \frac{3 c_2 - e^2}{720}, \quad ch(T^*X_{CY}) = 5 - c_2 + \frac{c_2^2 - 2 e_4}{12} + \frac{-c_3 - c_2 \cdot (-c_4)}{24}$$

so that one gets finally the relation

$$\frac{e}{24} = h^{41} - h^{31} + h^{21} - h^{11}.$$

(9)

Taken together with the obvious relation

$$\frac{e}{2} = h^{11} - h^{41} + 2(h^{31} - h^{21}) + h^{22} - h^{32}$$

(10)

one can now express the cohomology completely in terms of the known numbers $e$, $\delta = h^{41} - h^{11}$ and the remaining unknowns $h^{31}$ and $h^{22}$

$$h^{31} = h^{21} + \delta + \frac{e}{24}, \quad h^{32} = h^{22} + \delta - \frac{5}{12} e.$$  

(11)

Also one has now for the 5-folds of the special form $CY^5 = X^4 \times \mu \cdot dP_9$ that (cf. appendix; $CY^3$ denotes the fibre of $X^4$)

$$h^{21} = h^{21}(X^4), \quad h^{22} = 10 h^{11}(CY^3) + 2 h^{21}(X^4) + h^{22}(X^4) + 1.$$  

(12)

Now let us come back to our example $CY^5 = dP_9 \times \mu \cdot K \times \mu \cdot dP_9$ of $e = 0$. This decomposition allows one to find for the Kähler classes $h^{11} = 10 - 1 + 3 + 10 - 1 = 21$. On the other hand one has (with $\# \text{def}dP_9 = 8$) for the complex deformations that $h^{41} = 8 + 3 + \# \text{def}K + 8 + 3$, so with $\# \text{def}K = 2 \cdot 10 \cdot 2 - (3 + 8 + 3) - 1 = 25$ one gets $h^{41} = 47$. Furthermore the decomposition shows (cf. the appendix) that $h^{21} = h^{21}(K) = \# \text{def}K - (\# \text{def}dP_9 - 1) = 2 \cdot 10 \cdot 2 - (3 + 8 + 3) - 1 - (8 - 1) = 18$ and $h^{22} = 10 \cdot 19 + 2 \cdot 18 + h^{22}(X^4) + 1$ which gives with $h^{22}(X^4) = 10 \cdot 10 + 2 h^{21}(X^4) + h^{11}(K) + 1 = 140$ that $h^{22} = 367$, so

$$\begin{array}{ccccccc}
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 18 & 18 & 0 & 47 & 393 \\
0 & 44 & 393 & 393 & 47 & 1 & 0
\end{array}$$

Similarly one can use any of the existing lists of Calabi-Yau 4-folds (for example with the STU-Calabi-Yau $P_{1,1,2,3,2,2}(24)$ as 3-fold fibre), go to the correspondingly reduced $X^4$ (model $X^4_4$ for the example just mentioned, cf. [11]) and describe a $CY^5$. 


4. The brane point of view

Just as one can interpret an F-theory vacuum either as a Calabi-Yau compactification of a formally twelve-dimensional theory or as a type IIB vacuum with varying dilaton and 7-branes one can use the alternative brane point of view for S-theory as well. This was studied especially for the $T^3$ part in [4]; let us recall this point of view first and then interpret our example that way.

So what is given according to this point of view is really an 8D vacuum configuration with varying moduli consistent with the U-duality group $SL(3,\mathbb{Z}) \times SL(2,\mathbb{Z})$. The relevant moduli space Eq. (1) leads as described to the idea of U-manifolds admitting a $T^3 \times T^2$ fibration. $T^2$ fibrations were studied in F-theory so let us focus on the five-dimensional piece of the moduli space parametrizing a three-torus of constant volume. This translates to a family of 5-branes that transform consistently with $SL(3,\mathbb{Z})$, living on $S^5$, where each individual member lives on a ’line’ (set of real codimension two) in the base $S^5$. (In addition the 5-branes are also wrapped around the $S^2$ part of the $B^5$ base.) Phrased differently (and making it comparable to the cosmic string viewpoint in F-theory) the solution may be viewed as a mapping of the base into the five-dimensional moduli space, which is locally $SL(3,\mathbb{R})/SO(3)$ and actually an orbifold from the identification under the action of $SL(3,\mathbb{Z})$ U-duality. The pullback of the orbifold singularities leads then to the 5-brane configuration wrapping the singular lines (compare the F-theory picture relating the degenerate elliptic curves with the 7-brane locations; here the $T^3$’s are expected to degenerate along the singular lines, which correspond to the one-dimensional compact part of the world volume of the 5-branes; but note that in the F-theoretic case with the $K3$ there are only parallel branes involved).

Let us describe the real 5-dimensional base space $B^5$ for the $T^3 \times T^2$ fibration of $CY^5 = X^4 \times_p dP_9$ more fully. It is as already remarked a $S^3$ fibration over $S^5$. This simply connected space has the one non-trivial Betti number $b_2 = 1 (= b_3)$. Because of $\pi_1(SO(4)) = \mathbb{Z}_2$ there exist actually only two possibilities for $B^5$, either the product $S^2 \times S^3$ or the twisted version. Furthermore, taking into account the possibility to represent the $S^3$ fibre itself as a $S^1$ fibration over $S^2$ (the Hopf fibration), the mentioned ambiguity for the $B^5$ space can, because of $\pi_1(SO(3)) = \mathbb{Z}_2$, already be read of from the 4-manifold consisting of the $S^3$ fibration over the base $S^5$. Note that, in the case of a complex structure for this 4-manifold, the ‘fibre-type’ ($\in \mathbb{Z}_2$) ambiguity of the Hirzebruch surface $F_n$ (being a $P^1_f = S^3_f$ fibration over the base $P^1 = S^5$) is described by $n$ being even/odd (note the deformation $F_2 \rightarrow F_0$).

The $T^2$ fibration over $S^5$ of the $CY^3$ fibre will be combined with the $T^2$ fibration over $P^1_8$ of the $dP^d_9$. The singular lines in the $S^3$ are now replaced by singular loci of real codimension two in the base $B^5$ (which itself is a $S^3$ fibration over $S^2$). Note that we get a 5-brane picture in total as the $T^2$ fibration part gives also 5-branes: namely 7-branes (cf. F-theory) compactified on the $T^2$ which brought us from 10 to 8 dimensions; note that here the 7-branes have their locus not on a $P^1$, compactifying 10 dimensions to 8 dimensions, but on the $P^1_8$, compactifying from 5 dimensions to 3 dimensions. Of course the relevant singular loci consist now in the singular lines of the three-fold with their parameters running over the $S^2_{\pm}$ base of $B^5$ on the one hand, and furthermore in the $S^2$ fibers over the 12 singular points (for the $dP^d_9$) on $S^2_\pm$. So one gets in both cases subspaces of real dimension three in $B^5$, i.e. the loci of the 5-branes are of real codimension 2 in the base.

Finally let us also consider the analogue of the (now not internal but spacetime-filling) 3-branes which have to be turned on for F-theory on a 4-fold. These are, as already mentioned in [3], (spacetime-filling) 4-branes in the case of S-theory on a 4-fold. Now our $CY^5$ is $CY^4$ fibered over $P^1$; so in this further compactification process the 4-branes wrap the $P^1$ and become (spacetime-filling) membranes in three dimensions. But as our $CY^4$ fibre had to be the reducible $CY^3 \times T^2$ of Euler number zero, the mentioned branes do not actually occur.

5. The spectrum

Let us now read of from the Hodge diamond some part of the spectrum of massless multiplets. We will do this by the same strategy which is used to get a corresponding part of the F-theory spectrum from information
about type II compactifications [2,11]. This uses that F-theory on $X \times T^2$ is type IIA on $X$ and furthermore that F-theory, being partly simply type IIB on the basis $B$ of the relevant Calabi-Yau space in question, shows a sensitivity on the Hodge numbers of $B$ (which is not seen any longer - after further compactification on $T^2$ - in the type IIA description). Now in our case here we will use the same procedure and relate S-theory after further B analysis on $B$ space.

Now, the searched for spectrum is that of S-theory on $CY^5$. This leads to $N = 1$ supersymmetry in three dimensions. Beside the $N = 1$ supergravity multiplet, which contains as its bosonic freedom the metric $g_{\mu\nu}$, there will be $S_1$ real $N = 1$ scalar multiplets plus $V_1$ real $N = 1$ vector multiplets. The on-shell degrees of freedom of each $N = 1$ scalar multiplet are given by one real scalar field plus one real Majorana spinor; the $N = 1$ vector multiplets in three dimensions contain, on-shell, one vector field plus one real Majorana spinor. Since in three dimensions a vector is dual to a scalar, there is a (supersymmetric) Poincare duality between the scalar and the vector multiplets.

To obtain the spectrum of S-theory on $CY^5$ we start with the consideration of the type IIB superstring in ten dimensions. Its massless bosonic fields are

$$g_{MN}, B_{MN}, \phi, \phi', A_{MN}, A_{MNP}^\gamma.$$ (13)

To obtain S-theory vacua we first have to compactify the type IIB superstring on a two-dimensional torus $T^2$ to eight dimensions. This leads to non-chiral eight-dimensional $N = 2$ supergravity (like the type IIA compactification on $T^2$). The only massless supermultiplet is the supergravity multiplet. From Eq. (13) it is easy to see that the eight-dimensional $N = 2$ supergravity multiplet contains the following massless bosonic fields:

$$g_{MN}, 7\phi, 6A_M, 3A_{MN}, A_{MNP}.$$ (14)

(Now the indices $M, N, \ldots$ run over $0, \ldots, 7$; note that a possible 4-form is dual to the 2-form coming from the 10D 4-form which (i.e. its field-strength) is self-dual.) The seven scalar fields parametrize the non-compact coset space Eq. (1).

At the next step we compactify this eight-dimensional theory down to three dimensions on the S-theory base space $B^5$ to obtain the base-sensitive part of the three-dimensional spectrum. One performs the harmonic analysis on $B^5$ deriving from Eq. (14) the following contributions $s_3$ and $e_3$ to the number of scalar and $U(1)$ vector fields (as $b_1 = 0$, $b_2 = b_3 = 1$):

$$s_3 = 7 + 6b_1 + 3b_2 + 1b_3 = 11, \quad e_3 = 6 + 3b_1 + b_2 = 7.$$ (15)

Now we consider M-theory on $CY^5$, which leads to $N = 2$ supergravity in one dimension. There are $V^M_1$ vector multiplets with one real physical scalar and one non-propagating vector (plus one non-propagating scalar). In addition we will have $S^M_1$ scalar multiplets with each one physical scalar field. The internal metric of $CY^5$ provides $h^{+1,1} + 2h^{+1}$ real scalars. The 11-dimensional field $A_{MNP}$ will contribute in addition $h^{1,1}$ $U(1)$ vectors plus $2h^{+1}$ scalars. So in total we derive:

$$V^M_1 = h^{1,1}, \quad S^M_1 = 2h^{2,1} + 2h^{4,1}.$$ (16)
This $M$-theory spectrum can be directly compared with the $S$-theory spectrum on $CY^5 \times T^2$. The three-dimensional massless spectrum of $S$-theory, denoted by $S_3$ and $V_3$ is related to the one dimensional $S$-theory spectrum as follows:

$$V_1^S = V_3 + 2, \quad S_1^S = S_3.$$  \hfill (17)

So with Eq. (16) the $S$-theory/M-theory duality in one dimension leads to the constraint $V_1^S + S_1^S = V_1^M + S_1^M$, and hence we derive

$$V_3 + S_3 = h^{1,1} + 2 h^{2,1} + 2 h^{4,1} - 2.$$  \hfill (18)

Including the base-sensitive part one gets

$$S_3 = 2 h^{2,1} + 2 h^{4,1} + s_3 - v_3 = 2 h^{2,1} + 2 h^{4,1} + 4, \quad V_3 = h^{1,1} + v_3 - s_3 - 2 = h^{1,1} - 6.$$  \hfill (19)

Note that the $V_3$ only counts the massless abelian vector fields; at special loci in the moduli space additional non-abelian gauge bosons together with charged matter fields are expected to become massless.

Acknowledgements

We like to thank A. Miemiec, R. Minasian, C. Vafa and E. Witten for useful discussions.

Appendix A. 5-folds as fibre products

For more details on many of the mentioned spaces cf. for example [11,10].

The surface

$$dP_9 = \begin{bmatrix} P^2 & 3 \\ P^1 & 1 \end{bmatrix}$$

has the one nontrivial Hodge number $h^{11} = 10$ and $8 = 10 \cdot 2 - (8 + 3) - 1$ complex deformations. Note that you can visualize the 10 classes on the one hand by viewing the surface $P^2$ blown up in the 9 intersection points of the two cubics (in this sense it is a generalization of the del Pezzo surfaces $dP_i$ for $i = 1, \ldots, 8$); on the other hand you can understand their appearance topologically in the elliptic fibration picture via the fact that an $S^1$ of the fibre moving between two vanishing points traces out an $S^2 = P^1$ (cf. [13]).

The Calabi-Yau three-fold

$$CY^{19,19} = \begin{bmatrix} P^2 & 3 & 0 \\ P^1 & 1 & 1 \\ P^2 & 0 & 3 \end{bmatrix} = dP_9 \times P^1, dP_9$$

has obviously $h^{11} = 10 + 10 - 1$ and so from $e = 0$ also $h^{21} = 19$, which you can also count directly as $8 + 3 + 8$ (as one can use the reparametrization freedom on the $P^1$ only once). Note that also $CY^{19,19} = (K3 \times T^2)/Z_2$ (here the first of the two $dP_9$-fibre factors is appearing ‘by reduction’ from the former

$$K3 = \begin{bmatrix} P^2 \\ P^1 \\ 2 \end{bmatrix}$$
factor, the second one is ‘emerging’ from the constant $T^2$-factor in the process of smoothing out the quotient.

The Calabi-Yau *four-fold*

$$
CY^4 = \begin{bmatrix}
p^2 & 3 & 0 \\
p^1 & 1 & 1 \\
p^1 & 0 & 2 \\
p^2 & 0 & 3 \\
\end{bmatrix} = dP_9 \times p_1 B
$$

with

$$
B := \begin{bmatrix}
p^1 & 1 \\
p^1 & 2 \\
p^2 & 3 \\
\end{bmatrix}
$$

has the following Hodge diamond (model A, cf. [11])

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>28</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>28</td>
<td>260</td>
<td>260</td>
<td>56</td>
<td>56</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>140</td>
<td>428</td>
<td>428</td>
<td>140</td>
<td>140</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

where also $h^{21} = h^{21}(B)$ (cf. for this numerically [11]; it can be seen also from the topological ‘tracing out’ argument above that no new classes appear: just as the $K3$ fibration of $B$ over $P^1$ shows how a $S^2$ moving over the base between two vanishing points traces out a $S^5$ for $h^{21}(B)$, the corresponding $T^2 \times K3$ fibration over $P^1$ of $CY^4 = dP_9 \times p_1 B$ shows that no new classes appear). Note also that $CY^4 = (T^2 \times CY^3)/Z_2$. There is an ‘alternative version’ of this space which uses instead of the

$$
CY^3 = \begin{bmatrix}
p^1 & 2 \\
p^1 & 2 \\
p^2 & 3 \\
\end{bmatrix},
$$

from which $B$ is derived (by quadratic base change in the one $P^1$), the well-known $CY^{3,243}$; this leads to a $CY^4$ with Hodge-diamond shown above as model A [11].

So one has in both cases

$$
h^{22} = 204 + 2 h^{21}. \quad (A.1)
$$

This relation which of course is not accidental has two explanations: a numerical one and a geometrical one. The latter will be of relevance for our understanding of $h^{22}$ of a $CY^5$.

Now first the numerical argument: one has

$$
h^{22} = e - 4 - 2(h^{11} - 2h^{21} + h^{31}) = \frac{2}{3} e + 12 + 2h^{21} \quad (A.2)
$$

(using $h^{11} - h^{21} + h^{31} = \frac{c}{e} - 8$ (cf. [15])) and $e = 12 \cdot 24 = 288$ shows the relation asserted above. Secondly the
geometrical interpretation makes visible the classes (of the relevant Hodge type) from the following four topological sources of 4-cycles:

\[ S^2 \times S^2; 
S^1 \times \text{3-cycle}: 2 h^{21}; 
4\text{-cycle } \times \text{point}: h^{22}(\mathcal{B}) = h^{11}(\mathcal{B}) = 3; 
\text{point } \times dP_9: 1. \]

So if we come now to the Calabi-Yau five-fold \( CY^5 = X^4 \times_P dP_9 \), we have again \( h^{21} = h^{21}(X^4) \) and in the case of \( X^4 = dP_9 \times_P K \) further \( h^{21} = h^{21}(K) \). Also we have from the geometrical arguments showing how \( h^{22} \) arises that (let \( CY^5 \) be the fibre of \( X^4 \) over \( P^4 \))

\[ h^{22} = 10 h^{11}(CY^5) + 2 h^{21}(X^4) + h^{22}(X^4) + 1. \]  (A.3)

References

Gauge theory correlators from non-critical string theory

S.S. Gubser 1, I.R. Klebanov 2, A.M. Polyakov 3

Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544, USA

Received 20 February 1998
Editor: M. Dine

Abstract

We suggest a means of obtaining certain Green’s functions in 3 + 1-dimensional \( \mathcal{N} = 4 \) supersymmetric Yang-Mills theory with a large number of colors via non-critical string theory. The non-critical string theory is related to critical string theory in anti-deSitter background. We introduce a boundary of the anti-deSitter space analogous to a cut-off on the Liouville coordinate of the two-dimensional string theory. Correlation functions of operators in the gauge theory are related to the dependence of the supergravity action on the boundary conditions. From the quadratic terms in supergravity we read off the anomalous dimensions. For operators that couple to massless string states it has been established through absorption calculations that the anomalous dimensions vanish, and we rederive this result. The operators that couple to massive string states at level \( n \) acquire anomalous dimensions that grow as \( 2(\alpha_s N^2)^{1/2} \) for large \( \alpha_s \) coupling. This is a new prediction about the strong coupling behavior of large \( N \) SYM theory. © 1998 Published by Elsevier Science B.V. All rights reserved.

1. Introduction

Relations between gauge fields and strings present an old, fascinating and unanswered question. The full answer to this question is of great importance for theoretical physics. It will provide us with a theory of quark confinement by explaining the dynamics of color-electric fluxes. On the other hand, it will perhaps uncover the true “gauge” degrees of freedom of the fundamental string theories, and therefore of gravity.

The Wilson loops of gauge theories satisfy the loop equations which translate the Schwinger-Dyson equations into variational equations on the loop space [1,2]. These equations should have a solution in the form of the sum over random surfaces bounded by the loop. These are the world surfaces of the color-electric fluxes. For the \( SU(N) \) Yang-Mills theory they are expected to carry the \( \alpha_s \) factor [3], \( N^X \), where \( X \) is the Euler character. Hence, in the large \( N \) limit where \( g_{YM}^2 N \) is kept fixed only the simplest topologies are relevant.

Until recently, the action for the “confining string” had not been known. In [4] was suggested that it must have a rather unusual structure. Let us describe this briefly. First of all, the world surface of the electric flux propagates in at least 5 dimensions.

\[
X^\mu(\sigma), \quad g_{ij}(\sigma) = e^{g(\sigma) S_{ij}}, \tag{1}
\]
where $X^\mu$ belong to 4-dimensional (Euclidean) space and $g_{ij}(\sigma)$ is the world sheet metric in the conformal gauge. The general form of the world sheet lagrangian compatible with the 4-dimensional symmetries is

$$\mathcal{L} = \frac{1}{2} (\partial_\sigma X^\mu)^2 + a^2(\sigma) \left( \partial_\tau X^\mu \right)^2 + \Phi(\sigma)^{(2)} R + \text{Ramond - Ramond backgrounds},$$

where $^{(2)}R$ is the world sheet curvature, $\Phi(\sigma)$ is the dilaton [5], while the field $\Sigma(\sigma) = a^2(\sigma)$ defines a variable string tension. In order to reproduce the zig-zag symmetry of the Wilson loop, the gauge fields must be located at a certain value $\sigma = \varphi$, such that $a(\varphi, \lambda = 0) = 0$. We will call this point "the horizon".

The background fields $\Phi(\sigma), a(\sigma)$ and others must be chosen to satisfy the conditions of conformal invariance on the world sheet [6]. After this is done, the relation between gauge fields and strings can be described as an isomorphism between the general Yang-Mills operators of the type

$$\int d^4x e^{ip \cdot x} \text{tr} \left( \nabla_{\alpha_1} \ldots F_{\mu_1 \nu_1} \ldots \nabla_{\alpha_n} \ldots F_{\mu_n \nu_n}(x) \right)$$

and the algebra of vertex operators of string theory, which have the form

$$V^{a_1 \ldots a_s}(p) = \int d^2\sigma \Psi_{\rho_1 \ldots \rho_{10}}^{a_1 \ldots a_s}(\varphi(\sigma)) \times e^{i p \cdot X(\sigma)} H_{a_1} X^{a_2} \ldots \partial_{i_1} X^{\alpha_i} \partial_{j_1} \varphi \ldots \partial_{i_n} \varphi,$$

where the wave functions $\Psi^{a_1 \ldots a_s}_{\rho_1 \ldots \rho_{10}}(\varphi)$ are again determined by the conformal invariance on the world sheet. The isomorphism mentioned above implies the coincidence of the correlation functions of these two sets of vertex operators.

Another, seemingly unrelated, development is connected with the Dirichlet brane [7] description of black 3-branes in [8–11]. The essential observation is that, on the one hand, the black branes are solitons which curve space [12] and, on the other hand, the world volume of $N$ parallel D-branes is described by supersymmetric $U(N)$ gauge theory with 16 supercharges [13]. A particularly interesting system is provided by the limit of a large number $N$ of coincident D3-branes [8–11], whose world volume is described by $\mathcal{N} = 4$ supersymmetric $U(N)$ gauge theory in $3 + 1$ dimensions. For large $g_{YM}^2 N$ the curvature of the classical geometry becomes small compared to the string scale [9], which allows for comparison of certain correlation functions between the supergravity and the gauge theory, with perfect agreement [9–11]. Corrections in powers of $\alpha'$ times the curvature on the string theory side correspond to corrections in powers of $(g_{YM}^2 N)^{-1/2}$ on the gauge theory side. The string loop corrections are suppressed by powers of $1/N^2$.

The vertex operators introduced in [9–11] describe the coupling of massless closed string fields to the world volume. For example, the vertex operator for the dilaton is

$$\int \frac{1}{4 g_{YM}^2} d^4 x e^{ip \cdot x} \text{tr} \left( \tau F_{\mu \nu} F^{\mu \nu}(x) + \ldots \right),$$

while that for the graviton polarized along the 3-branes is

$$\int d^4x e^{ip \cdot x} T^{\mu \nu}(x),$$

where $T^{\mu \nu}$ is the stress tensor. The low energy absorption cross-sections are related to the 2-point functions of the vertex operators, and turn out to be in complete agreement with conformal invariance and supersymmetric non-renormalization theorems [11]. An earlier calculation of the entropy as a function of temperature for $N$ coincident D3-branes [8] exhibits a dependence expected of a field theory with $O(N^2)$ massless fields, and turns out to be $3/4$ of the free field answer. This is not a discrepancy since the free field result is valid for small $g_{YM}^2 N$, while the result of [8] is applicable as $g_{YM}^2 N \to \infty$. We now regard this result as a non-trivial prediction of supergravity concerning the strong coupling behavior of $\mathcal{N} = 4$ supersymmetric gauge theory at large $N$ and finite temperature.

The non-critical string and the D-brane approaches to $3 + 1$ dimensional gauge theory have
been synthesized in [14] by rescaling the 3-brane metric and taking the limit in which it has conformal symmetry, being the direct product $AdS_5 \times S^5$. This is exactly the confining string ansatz [4] with
\begin{equation}
 a(\varphi) = e^{\varphi/R},
\end{equation}
corresponding to the case of constant negative curvature of order $1/R^2$. The horizon is located at $\varphi_c = -\infty$. The Liouville field is thus related to the radial coordinate of the space transverse to the 3-brane. The extra $S^5$ part of the metric is associated with the 6 scalars and the $SU(4)$ R-symmetry present in the $\mathcal{N} = 4$ supersymmetric gauge theory.

In the present paper we make the next step and show how the masses of excited states of the "confining string" are related to the anomalous dimensions of the SYM theory. Hopefully this analysis will help future explorations of asymptotically free gauge theories needed for quark confinement.

We will suggest a potentially very rich and detailed means of analyzing the throat-brane correspondence: we propose an identification of the generating function of the Green’s functions of the superconformal world-volume theory and the supergravity action in the near horizon background geometry. We will find it necessary to introduce a boundary of the $AdS_5$ space near the place where the throat turns into the asymptotically flat space. Thus, the anti-deSitter coordinate $\varphi$ is defined on a half-line ($-\infty,0$], similarly to the Liouville coordinate of the 2-dimensional string theory [15,16]. The correlation functions are specified by the dependence of the action on the boundary conditions, again in analogy with the $c = 1$ case. One new prediction that we will be able to extract this way is for the anomalous dimensions of the gauge theory operators that correspond to massive string states. For a state at level $n$ we find that, for large $g_{YM}^2 N$, the anomalous dimension grows as $2\sqrt{n}/(2 g_{YM}^2 N)^{1/4}$.

\footnote{The papers that put an early emphasis on the anti-deSitter nature of the near-horizon region of certain brane configurations, and its relation with string and M-theory, are [17,18]. Other ideas on the relation between branes and AdS supergravity were recently pursued in [19–21].}

2. Green’s functions from the supergravity action

The geometry of a large number $N$ of coincident D3-branes is
\begin{equation}
 ds^2 = \left( \frac{1 + R^4}{r^4} \right)^{1/2} \left( -dt^2 + dx^2 \right) + \left( \frac{1 + R^4}{r^4} \right)^{-1/2} \left( dr^2 + r^2 d\Omega_5^2 \right).
\end{equation}

The parameter $R$, where
\begin{equation}
 R = \frac{N}{2 \pi^2 T_3}, \quad T_3 = \frac{\sqrt{\pi}}{\kappa}
\end{equation}
is the only length scale involved in all of what we will say. $T_3$ is the tension of a single D3-brane, and $\kappa$ is the ten-dimensional gravitational coupling. The near-horizon geometry of $N$ D3-branes is $AdS_5 \times S^5$, as one can see most easily by defining the radial coordinate $z = R^2/r$. Then
\begin{equation}
 ds^2 = \frac{R^2}{z^2} \left( -dt^2 + dx^2 + dz^2 \right) + R^2 d\Omega_5^2.
\end{equation}
The relation to the coordinate $\varphi$ used in the previous section is
\begin{equation}
 z = R e^{-\varphi/R}. \tag{11}
\end{equation}

\footnote{Note that the limit $z \to 0$ is far from the brane. Of course, for $z \ll R$ the $AdS$ form (10) gets modified, and for $z \ll R$ one obtains flat ten-dimensional Minkowski space. We will freely use phrases like "far from the brane" and "near the brane" to describe regimes of small $z$ and large $\varphi$, despite the fact that the geometry is geodesically complete and nonsingular.}

The basic idea is to identify the generating functional of connected Green’s functions in the gauge theory with the minimum of the supergravity action, subject to some boundary conditions at $z = R$ and $z = \infty$:
\begin{equation}
 W \left[ g_{\mu\nu}(x^4) \right] = K \left[ g_{\mu\nu}(x^4) \right] = S \left[ g_{\mu\nu}(x^4, z) \right]. \tag{12}
\end{equation}

$W$ generates the connected Green’s functions of the gauge theory; $S$ is the supergravity action on the
AdS space; while $K$ is the minimum of $S$ subject to the boundary conditions. We have kept only the metric $g_{\mu\nu}(x^k)$ of the world-volume as an explicit argument of $W$. The boundary conditions subject to which the supergravity action $S$ is minimized are

$$ds^2 = \frac{R^2}{z^2} \left( g_{\mu\nu}(x^k) dx^\mu dx^\nu + dz^2 \right) + O(1) \quad \text{as} \quad z \to R.$$  \hspace{1cm} (13)

All fluctuations have to vanish as $z \to \infty$.

A few refinements of the identification (12) are worth commenting on. First, it is the generator of connected Green’s functions which appears on the left hand side because the supergravity action on the right hand side is expected to follow the cluster decomposition principle. Second, since classical supergravity is reliable only for a large number $N$ of coincident branes, (12) can only be expected to capture the leading large $N$ behavior. Corrections in $1/N$ should be obtained as loop effects when one replaces the classical action $S$ with an effective action $T$. This is sensible since the dimensionless expansion parameter $\kappa^2/R^8 \sim 1/N^2$. We also note that, since $(\alpha')^2/R^4 \sim (g_{YM}^2 N)^{-1}$, the string theoretic $\alpha'$ corrections to the supergravity action translate into gauge theory corrections proportional to inverse powers of $g_{YM}^2 N$. Finally, the fact that there is no covariant action for type IIB supergravity does not especially concern us: to obtain $n$-point Green’s functions one is actually considering the $n$th variation of the action, which for $n > 0$ can be regarded as the $(n-1)$th variation of the covariant equations of motion.

In Section 2.2 we will compute two-point functions of massless vertex operators from (12), compare them with the absorption calculations in [9–11] and find exact agreement. However it is instructive first to examine boundary conditions.

2.1. Preliminary: symmetries and boundary conditions

As a preliminary it is useful to examine the appropriate boundary conditions and how they relate to the conformal symmetry. In this discussion we follow the work of Brown and Henneaux [22]. In the consideration of geometries which are asymptotically anti-de Sitter, one would like to have a realization of the conformal group on the asymptotic form of the metric. Restrictive or less restrictive boundary conditions at small $z$ (far from the brane) corresponds, as Brown and Henneaux point out in the case of $AdS_3$, to smaller or larger asymptotic symmetry groups. On an $AdS_{d+1}$ space,

$$ds^2 = G_{mn} dx^m dx^n = \frac{R^2}{z^2} (- dt^2 + dx^2 + dz^2)$$  \hspace{1cm} (14)

where now $x$ is $d - 1$ dimensional, the boundary conditions which give the conformal group as the group of asymptotic symmetries are

$$\delta G_{\mu\nu} = O(1), \quad \delta G_{\mu\nu} = O(z), \quad \delta G_{zz} = O(1).$$  \hspace{1cm} (15)

Our convention is to let indices $m,n$ run from $0$ to $d$ (that is, over the full $AdS_{d+1}$ space) while $\mu,\nu$ run only from $0$ to $d - 1$ (i.e. excluding $z = x^d$). Diffeomorphisms which preserve (15) are specified by a vector $\xi^m$ which for small $z$ must have the form

$$\xi^k = \xi^k - \frac{z^2}{d} \eta^{\mu\nu} \partial_{\mu}(\xi^\nu) + O(z^4)$$

$$\xi^z = \frac{z}{d} \xi^z + O(z^3).$$  \hspace{1cm} (16)

Here $\xi^m$ is allowed to depend on $t$ and $x$ but not $r$. (16) specifies only the asymptotic form of $\xi^m$ at large $r$, in terms of this new vector $\xi^m$.

Now the condition that the variation

$$\delta G_{mn} = \partial_m \xi^n + G_{mn} \partial_n \xi^k + G_{kn} \partial_n \xi^m$$  \hspace{1cm} (17)

be of the allowed size specified in (15) is equivalent to

$$\xi_{\mu,\nu} + \xi_{\nu,\mu} = \frac{2}{d} \xi^k$$  \hspace{1cm} (18)

where now we are lowering indices on $\xi^\mu$ with the flat space Minkowski metric $\eta_{\mu\nu}$. Since (18) is the conformal Killing equation in $d$ dimensions ($d = 4$ for the 3-brane), we see that we indeed recover precisely the conformal group from the set of permissible $\xi^m$. 

The spirit of [22] is to determine the central charge of an $AdS_3$ configuration by considering the commutator of deformations corresponding to Virasoro generators $L_m$ and $L_{-m}$. This method is not applicable to higher dimensional cases because the conformal group becomes finite, and there is apparently no way to read off a Schwinger term from commutators of conformal transformations. Nevertheless, the notion of central charge can be given meaning in higher dimensional conformal field theories, either via a curved space conformal anomaly (also called the gravitational anomaly) or as the normalization of the two-point function of the stress energy tensor [23]. We shall see in Section 2.2 that a calculation reminiscent of absorption probabilities allows us to read off the two point function of stress-energy tensors in $\mathcal{N} = 4$ super-Yang-Mills, and with it the central charge.

### 2.2. The two point functions

First we consider the case of a minimally coupled massless scalar propagating in the anti-de Sitter near-horizon geometry (one example of such a scalar is the dilaton $\phi$ [9]). As a further simplification we assume for now that $\phi$ is in the $s$-wave (that is, there is no variation over $S^3$). Then the action becomes

$$
S = \frac{1}{2\kappa^2} \int d^{10}x \sqrt{G} \left[ \frac{1}{2} G^{MN} \partial_M \phi \partial_N \phi \right] = \frac{\pi^3 R^8}{4\kappa^2} \int d^4x \int_R^{\infty} \frac{dz}{z^3} \left[ (\partial_z \phi)^2 + \eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi \right].
$$

Note that in (19), as well as in all the following equations, we take $\kappa$ to be the ten-dimensional gravitational constant. The equations of motion resulting from the variation of $S$ are

$$
\left[ z^4 \partial_z \frac{1}{z^4} \partial_z + \eta^{\mu\nu} \partial_\mu \partial_\nu \right] \phi = 0.
$$

A complete set of normalizable solutions is

$$
\phi_s(x^\prime) = \lambda_s e^{ik\cdot x^\prime} \tilde{\phi}_s(z),
$$

where $\tilde{\phi}_s(z) = \frac{z^2 K_2(kz)}{R^2 K_2(kR)}$, $k^2 = k^2 - \omega^2$.

We have chosen the modified Bessel function $K_2(kz)$ rather than $I_2(kz)$ because the functions $K_2(kz)$ fall off exponentially for large $z$, while the functions $I_2(kz)$ grow exponentially. In other words, the requirement of regularity at the horizon (far down the throat) tells us which solution to keep. A connection of this choice with the absorption calculations of [9] is provided by the fact that, for time-like momenta, this is the incoming wave which corresponds to absorption from the small $z$ region. $\lambda_s$ is a coupling constant, and the normalization factor has been chosen so that $\tilde{\phi}_s = 1$ for $z = R$.

Let us consider a coupling

$$
S_\text{int} = \int d^4x \phi(x^4) \mathcal{O}(x^4)
$$

in the world-volume theory. If $\phi$ is the dilaton then according to [9] one would have $\mathcal{O} = \frac{1}{16\pi^2} F^2 + \ldots$. Then the analogue of (12) is the claim that

$$
W[\phi(x^4)] = K[\phi(x^4)] = S[\phi(x^4, z)]
$$

where $\phi(x^4, z)$ is the unique solution of the equations of motion with $\phi(x^4, z) \to \phi(x^4)$ as $z \to R$. Note that the existence and uniqueness of $\phi$ are guaranteed because the equation of motion is just the Laplace equation on the curved space. (One could in fact compactify $x^4$ on very large $T^4$ and impose the boundary condition $\phi(x^4, z) = \phi(x^4)$ at $z = R$. Then the determination of $\phi(x^4, z)$ is just the Dirichlet problem for the laplacian on a compact manifold with boundary).

Analogously to the work of [15] on the $c = 1$ matrix model, we can obtain the quadratic part of $K[\phi(x^4)]$ as a pure boundary term through integration by parts,

$$
K[\phi(x^4)] = \frac{\pi^3 R^8}{4\kappa^2} \int d^4x \int_R^{\infty} \frac{dz}{z^3}
$$

$$
\times \left[ \frac{1}{z^4} \partial_z \frac{1}{z^4} \partial_z + \eta^{\mu\nu} \partial_\mu \partial_\nu \right] \phi
$$

$$
+ \frac{1}{z^4} \partial_z \left( \frac{1}{z^4} \partial_z \phi \right)
$$

$$
= \frac{1}{2} \int d^4k d^4q \lambda_s \lambda_q (2\pi)^4
$$

$$
\times \delta^4(k + q) \frac{N^2}{16\pi^2} \mathcal{O}.\quad (24)
$$
where we have expanded
\[ \phi(x) = \int d^4k \lambda_k e^{ik \cdot x}. \]

The “flux factor” \( \mathcal{F} \) (so named because of its resemblance to the particle number flux in a scattering calculation) is
\[ \mathcal{F} = \left[ \frac{1}{z^3} \delta_z \Delta \phi_k \right]_{z = R}. \tag{25} \]

In (24) we have suppressed the boundary terms in the \( x^4 \) directions—again, one can consider these compactified on very large \( T^4 \) so that there is no boundary. We have also used (9) to simplify the prefactor. Finally, we have cut off the integral at the boundary. We have also used 9 to simplify the compactified on very large \( T^4 \) so that there is no boundary.

To calculate the two-point function of \( \mathcal{O} \) in the world-volume theory, we differentiate \( K \) twice with respect to the coupling constants \( \lambda \) :
\[ \langle \mathcal{O}(k) \mathcal{O}(q) \rangle = \int d^4x d^4y e^{ik \cdot x + iq \cdot y} \langle \mathcal{O}(x) \mathcal{O}(y) \rangle = \frac{\delta^2 K}{\partial \lambda_k \partial \lambda_q} \]
\[ = (2\pi)^4 \delta^4(k + q) \frac{N^2}{16\pi^2} \mathcal{F} \]
\[ = -(2\pi)^4 \delta^4(k + q) \frac{N^2}{64\pi^2} k^4 \ln(k^2 R^2) \]
\[ + \text{(analytic in } k^2) \] \tag{26}

where now the flux factor has been evaluated as
\[ \mathcal{F} = \left[ \frac{1}{z^3} \delta_z \Delta \phi_k \right]_{z = R} = \left[ \frac{1}{z^3} \delta_z \Delta \phi_k \right]_{z = R} = -\frac{1}{4} k^4 \ln(k^2 R^2) + \text{(analytic in } k^2). \tag{27} \]

Fourier transforming back to position space, we find
\[ \langle \mathcal{O}(x) \mathcal{O}(y) \rangle \sim \frac{N^2}{|x - y|^2}. \tag{28} \]

This is consistent with the free field result for small

\[ g_{YM}^2 N. \text{ Remarkably, supergravity tells us that this formula continues to hold as } g_{YM}^2 N \to \infty. \]

Another interesting application of this analysis is to the two-point function of the stress tensor, which with the normalization conventions of [11] is
\[ \langle T_{\mu \nu}(x) T_{\rho \sigma}(0) \rangle = \frac{c}{48\pi^2} X_{\alpha \beta \gamma} \left( \frac{1}{x^4} \right), \tag{29} \]

where the central charge (the conformal anomaly) is
\[ c = N^2/4 \]
and
\[ X_{\alpha \beta \gamma} = 2 \Box \delta_{\alpha \beta} \eta_{\gamma \delta} - 3 \Box (\eta_{\gamma \delta} \delta_{\alpha \beta} + \eta_{\alpha \beta} \delta_{\gamma \delta}) \]
\[ - 4 \delta_{\gamma \delta} \eta_{\alpha \beta} - 2 \Box (\delta_{\alpha \beta} \eta_{\gamma \delta} + \delta_{\gamma \delta} \eta_{\alpha \beta} + \delta_{\delta \gamma} \eta_{\alpha \beta} + \delta_{\gamma \delta} \eta_{\alpha \beta}). \tag{30} \]

For metric perturbations \( g_{\mu \nu} = \eta_{\mu \nu} + h_{\mu \nu} \) around flat space, the coupling of \( h_{\mu \nu} \) at linear order is
\[ S_{lin} = \int d^4 x \frac{1}{2} h^{\mu \nu} T_{\mu \nu}. \tag{31} \]

Furthermore, at quadratic order the supergravity action for a graviton polarized along the brane, \( h_{\alpha \beta}(k) \), is exactly the minimal scalar action, provided the momentum \( k \) is orthogonal to the \( xy \) plane. We can therefore carry over the result (26) to obtain
\[ \langle T_{\alpha \beta}(k) T_{\gamma \delta}(q) \rangle = -(2\pi)^4 \delta^4(k + q) \frac{N^2}{64\pi^2} k^4 \ln(k^2 R^2) \]
\[ + \text{(analytic in } k^2). \tag{32} \]

which upon Fourier transform can be compared with (29) to give \( c = N^2/4 \). In view of the conformal symmetry of both the supergravity and the gauge theory, the evaluation of this one component is a sufficient test.

The conspiracy of overall factors to give the correct normalization of (32) clearly has the same origin as the successful prediction of the minimal scalar \( s \)-wave absorption cross-section [9–11]. The absorption cross-section is, up to a constant of proportionality, the imaginary part of (32). In [9] the absorption cross-section was calculated in supergravity using propagation of a scalar field in the entire 3-brane metric, including the asymptotic region far from the brane. Here we have, in effect, replaced communication of the throat region with the asymptotic region by a boundary condition at one end of the throat. The physics of this is clear: signals com-
ing from the asymptotic region excite the part of the throat near $z = R$. Propagation of these excitations into the throat can then be treated just in the anti-de-Sitter approximation. Thus, to extract physics from anti-de-Sitter space we introduce a boundary at $z = R$ and take careful account of the boundary terms that contain the dynamical information.

It now seems clear how to proceed to three-point functions: on the supergravity side one must expand to third order in the perturbing fields, including in functions: on the supergravity side one must expand and take careful account of the boundary terms that anti-deSitter approximation. Thus, to extract physics from the throat near $z$ into the throat can then be treated just in the anti-deSitter space we introduce a boundary at $z = R$.

The calculation is still simple in concept the classical particular the three point vertices. At higher orders to third order in the perturbing fields, including in functions: on the supergravity side one must expand contain the dynamical information.

One extension of the present work is to consider what fields couple to the other operators in the $N=4$ supercurrent multiplet. The structure of the multiplet (which includes the supercurrents, the $SU(4)$ $R$-currents, and four spin 1/2 and one scalar field) suggests a coupling to the fields of gauged $N=4$ supergravity. The question then becomes how these fields are embedded in $N = 8$ supergravity. We leave these technical issues for the future, but with the expectation that they are "bound to work" based on supersymmetry.

The main lesson we have extracted so far is that, for certain operators that couple to the massless string states, the anomalous dimensions vanish. We expect this to hold for all vertex operators that couple to the fields of supergravity. This may be the complete set of operators that are protected by supersymmetry. As we will see in the next section, other operators acquire anomalous dimensions that grow for large 't Hooft coupling.

3. Massive string states and anomalous dimensions

Before we proceed to the massive string states, a useful preliminary is to discuss the higher partial waves of a minimally coupled massless scalar. The action in five dimensions (with Lorentzian signature), the equations of motion, and the solutions are

$$ S = \frac{\pi^3 R^8}{4\kappa^2} \int d^4 x \int_0^\infty \frac{dz}{z^3} \times \left( (\partial_\phi \phi)^2 + (\bar{\partial}_\phi \phi)^2 + \frac{\ell (\ell + 4)}{z^2} \phi^2 \right) $$

$$ \left[ z^2 \eta_{\mu \nu} \frac{1}{z^2} \partial^\mu \phi + \bar{\eta}^\mu \bar{\partial}_\mu \phi - \frac{\ell (\ell + 4)}{z^2} \right] \phi = 0 $$

where $\bar{\phi}_k(z) = e^{ik \cdot \bar{R} k R}$.

We have chosen the normalization such that $\bar{\phi}_k(z) = 1$ at $z = R$. The flux factor is evaluated by expanding

$$ K_{\ell+2}(k z) = 2^{\ell+1} T(\ell + 2)(k z)^{-(\ell+2)} $$

$$ \times \left\{ 1 + \ldots + \frac{(-1)^\ell}{2^{2\ell+3} (\ell + 1)! (\ell + 2)!} \right\} $$

$$ \times (k z)^{(\ell+2)} \ln k z + \ldots $$.  

where in parenthesis we exhibit the leading non-analytic term. We find

$$ \mathcal{F} = \left[ \frac{1}{z^3} \partial_z \ln (\bar{\phi}_k) \right]_{z=R} $$

$$ = \frac{(-1)^\ell}{2^{2\ell+2} [(\ell + 1)!]^2} k^{4+2\ell} R^{2\ell} \ln k R $$.  

As before, we have neglected terms containing analytic powers of $k$ and focused on the leading nonanalytic term. This formula indicates that the operator that couples to $\ell$th partial wave has dimension $4 + \ell$. In [9] it was shown that such operators with the $SO(6)$ quantum numbers of the $\ell$th partial wave have the form

$$ \int d^4 x e^{ik \cdot \phi} \left\{ (X^{(i)} \ldots X^{(r)}) + \ldots \right\} F_{\mu \nu}^{\ell} (x) $$

$$ \text{tr} \left( X^{(i)} \ldots X^{(r)} + \ldots \right) F_{\mu \nu}^{\ell} (x) $$.
where in parenthesis we have a traceless symmetric tensor of $SO(6)$. Thus, supergravity predicts that their non-perturbative dimensions equal their bare dimensions.

Now let us consider massive string states. Our goal is to use supergravity to calculate the anomalous dimensions of the gauge theory operators that couple to them. To simplify the discussion, let us focus on excited string states which are spacetime scalars of mass $m$. The propagation equation for such a field in the background of the 3-brane geometry is

$$
\left[ \frac{d^2}{dr^2} + \frac{5}{r} \frac{d}{dr} - k^2 \left( 1 + \frac{R^4}{r^4} \right) \right] \bar{\phi}_k = 0. 
$$

(39)

For the state at excitation level $n$,

$$
m^2 = \frac{4n}{\alpha'}. 
$$

In the throat region, $z \gg R$, (39) simplifies to

$$
\left[ \frac{d^2}{dz^2} - \frac{3}{z} \frac{d}{dz} - k^2 \frac{m^2 R^2}{z^2} \right] \bar{\phi}_k = 0.
$$

(40)

Note that a massive particle with small energy $\omega \ll m$, which would be far off shell in the asymptotic region $z \ll R$, can nevertheless propagate in the throat region (i.e., it is described by an oscillatory wave function).

Eq. (40) is identical to the equation encountered in the analysis of higher partial waves, except the effective angular momentum is not in general an integer: in the centrifugal barrier term $\ell(\ell + 4)$ is replaced by $(mR)^2$. Analysis of the choice of wave function goes through as before, with $\ell + 2$ replaced by $\nu$, where

$$
\nu = \sqrt{4 + (mR)^2}. 
$$

(41)

In other words, the wave function falling off exponentially for large $z$ and normalized to 1 at $z = R$ is

$$
\bar{\phi}_k(z) = \frac{z^2 K_s(kz)}{R^2 K_s(kR)}. 
$$

(42)

Now we recall that

$$
K_s = \frac{\pi}{2 \sin(\pi \nu)}(I_{-\nu} - I_\nu),
$$

$$
I_\nu(z) = \left( \frac{z}{2} \right)^\nu \sum_{k=0}^\infty \frac{(z/2)^{2k}}{k! \Gamma(k + \nu + 1)}.
$$

Thus,

$$
K_s(kz) = 2^{-\nu-1}(\nu)(kz)^{-\nu} \times \left( 1 + \ldots - \left( \frac{z}{2} \right)^{2v} \frac{\Gamma(1 - \nu)}{\Gamma(1 + \nu)} + \ldots \right),
$$

(43)

where in parenthesis we have exhibited the leading non-analytic term. Calculating the flux factor as before, we find that the leading non-analytic term of the 2-point function is

$$
\langle \phi(k) \phi(q) \rangle = (2\pi)^d \delta^d(k + q) \times \frac{N^2}{8\pi^2} \frac{\Gamma(1 - \nu)}{\Gamma(1 + \nu)} \left( \frac{kR}{2} \right)^{2\nu} R^{-4}.
$$

(44)

This implies that the dimension of the corresponding SYM operator is equal to

$$
\Delta = 2 + \nu = 2 + \sqrt{4 + (mR)^2}. 
$$

(45)

Now, let us note that

$$
R^2 = 2Ng_\gamma^2 \ell(\alpha')^2,
$$

which implies

$$
(mR)^2 = 4ng_\gamma \sqrt{2N}. 
$$

Using (41) we find that the spectrum of dimensions

---

These solutions are reminiscent of the loop correlators calculated for $c \leq 1$ matrix models in [24].
for operators that couple to massive string states is, for large $g_{YM} \sqrt{N}$,

$$h_n \approx 2 \left( ng_{YM} \sqrt{2N} \right)^{1/2} . \quad (46)$$

Eq. (46) is a new non-trivial prediction of the string theoretic approach to large $N$ gauge theory. 7

We conclude that, for large 't Hooft coupling, the anomalous dimensions of the vertex operators corresponding to massive string states grow without bound. By contrast, the vertex operators that couple to the massless string states do not acquire any anomalous dimensions. This has been checked explicitly for gravitons, dilatons and RR scalars 9±11 , and we believe this to be a general statement. 8

Thus, there are several $SO(6)$ towers of operators that do not acquire anomalous dimensions, such as the dilaton tower (38). The absence of the anomalous dimensions is probably due to the fact that they are protected by SUSY. The rest of the operators are not protected and can receive arbitrarily large anomalous dimensions. While we cannot yet write down the explicit form of these operators in the gauge theory, it seems likely that they are the conventional local operators, such as (3). Indeed, the coupling of a highly excited string state to the world volume may be guessed on physical grounds. Since a string in a D3-brane is a path of electric flux, it is natural to assume that a string state couples to a Wilson loop $\mathcal{O}_{\exp iE A}$. Expansion of a small loop in powers of $F$ yields the local polynomial operators.

There is one potential problem with our treatment of massive states. The minimal linear equation (39) where higher derivative terms are absent may be true only for a particular field definition (otherwise corrections in positive powers of $\alpha' \nabla^2$ will be present in the equation). Therefore, it is possible that there are energy dependent leg factors relating the operators $\mathcal{O}$ in (44) and the gauge theory operators of the form (3). We hope that these leg factors do not change our conclusion about the anomalous dimensions. However, to completely settle this issue we need to either find an exact sigma-model which incorporates all $\alpha'$ corrections or to calculate the 3-point functions of massive vertex operators.

4. Conclusions

There are many unanswered questions that we have left for the future. So far, we have considered the limit of large 't Hooft coupling, since we used the one-loop sigma-model calculations for all operators involved. If this coupling is not large, then we have to treat the world sheet theory as an exact conformal field theory (we stress, once again, that the string loop corrections are $\sim 1/N^2$ and, therefore, vanish in the large $N$ limit). This conformal field theory is a sigma-model on a hyperboloid. It is plausible that, in addition to the global $O(2,4)$ symmetry, this theory possesses the $O(2,4)$ Kac-Moody algebra. If this is the case, then the sigma model is tractable with standard methods of conformal field theory.

Throughout this work we detected many formal similarities of our approach with that used in $c \leq 1$ matrix models. These models may be viewed as early examples of gauge theory – non-critical string correspondence, with the large $N$ matrix models playing the role of gauge theories. Clearly, a deeper understanding of the connection between the present work and the $c \leq 1$ matrix models is desirable.

Acknowledgements

We would like to thank C. Callan and the participants of the ‘‘Dualities in String Theory’’ program at ITP, Santa Barbara, for useful discussions. I.R.K. is grateful to ITP for hospitality during the completion of this paper. This research was supported in part by the National Science Foundation under Grants No. PHY94-07194 and PHY96-00258, by the Department of Energy under Grant No. DE-FG02-91ER40671, and by the James S. McDonnell Foundation under Grant No. 91-48. S.S.G. also thanks the Hertz Foundation for its support.

---

1 We do not expect this equation to be valid for arbitrarily large $n$, because application of linearized local effective actions to arbitrarily excited string states is questionable. However, we should be able to trust our approach for moderately excited states.

2 A more general set of such massless fields is contained in the supermultiplet of AdS gauge fields, whose boundary couplings were recently studied in [25].
References

Supersymmetry breaking with periodic potentials

Gerald Dunne\textsuperscript{a,1}, Jake Mannix\textsuperscript{b,2,3}

\textsuperscript{a} Physics Department, University of Connecticut, Storrs, CT 06269, USA
\textsuperscript{b} Physics Department, University of California, Santa Cruz, CA 93106, USA

Received 14 October 1997
Editor: M. Dine

Abstract

We discuss supersymmetry breaking in some supersymmetric quantum mechanical models with periodic potentials. The sensitivity to the parameters appearing in the superpotential is more acute than in conventional nonperiodic models. We present some simple elliptic models to illustrate these points. © 1998 Elsevier Science B.V. All rights reserved.

Supersymmetry (SUSY) breaking is an outstanding problem that is directly relevant for the application of SUSY to particle physics [1]. Certain aspects of this symmetry breaking mechanism may be studied using the simplest SUSY model, that of SUSY quantum mechanics [1]. Typically, one considers models with discrete spectra [1], while the extension to spectra with both bound and continuum states is relatively straightforward [2]. Recently, models with periodic potentials, which therefore have band spectra, have been considered [3]. The main new feature is that it is possible for the periodic isospectral bosonic and fermionic potentials to have exactly the same spectrum, including zero modes. Thus it is possible to have models with unbroken SUSY but for which the bosonic and fermionic hamiltonians have exactly identical spectra [3]. This is in contrast to the usual (nonperiodic and fast decaying) case for which at most one potential of an isospectral pair can have a zero mode. In this paper, we consider the breaking of SUSY in models with periodic superpotentials, and the sensitivity of these models to the parameters appearing in the superpotential.

SUSY quantum mechanics on the real line can be summarized as follows [1]. The bosonic and fermionic Hamiltonians $H_\pm$ correspond to an isospectral pair of potentials $V_\pm(x)$ defined in terms of the “superpotential” $W(x)$ as

\begin{equation}
V_\pm(x) = W^2(x) \pm W'(x)
\end{equation}

The Hamiltonians may be factorized into products of hermitean conjugate operators as

\begin{align}
H_+ &= \left[ \frac{d}{dx} + W(x) \right] \left[ -\frac{d}{dx} + W(x) \right],
\end{align}

\begin{align}
H_- &= \left[ -\frac{d}{dx} + W(x) \right] \left[ \frac{d}{dx} + W(x) \right].
\end{align}

\textsuperscript{1} E-mail: dunne@hep.phys.uconn.edu.
\textsuperscript{2} E-mail: jake@dirac.ucsc.edu.
\textsuperscript{3} Supported through the National Science Foundation R.E.U. (Research Experience for Undergraduates) program at the University of Connecticut.

0370-2693/98/$19.00$ © 1998 Elsevier Science B.V. All rights reserved.
PII: S0370-2693(98)00373-6
which indicates that $H_p$ are formally positive operators. The factorization (2) also implies that $V_{\pm}$ have (almost) the same spectrum because there is a one-to-one mapping between the energy eigenstates $\psi_\pm^\pm$:

$$\psi_\pm^\pm = \frac{1}{\sqrt{E}} \left( \frac{d}{dx} + W(x) \right) \psi_\pm^-;$$

$$\psi_\pm^- = \frac{1}{\sqrt{E}} \left( - \frac{d}{dx} + W(x) \right) \psi_\pm^+.$$  \hspace{1cm} (3)

This mapping between states does not apply to the "zero modes" (eigenstates with $E = 0$), which due to the positivity of $H_p$, are the lowest possible states in the spectrum. From (2) it is easy to see that the Schrödinger equation $[-\partial_x^2 + V_{\pm}(x)]\psi_\pm^\pm = E\psi_\pm^\pm$ has zero modes

$$\psi_\pm^\pm(x) = e^{\pm \int W}$$  \hspace{1cm} (4)

provided these functions $\psi_\pm^\pm$ belong to the Hilbert space. SUSY is unbroken if at least one of the $\psi_\pm^\pm$ is a true zero mode. Otherwise, SUSY is said to be broken. In the broken SUSY case there are no zero modes and so the spectra of $V_{\pm}$ are identical [due to the mapping (3)].

The Witten index, $\Delta = \text{tr}(-1)^F = n_+ - n_-$, counts the difference between the number ($n_+$) of bosonic zero modes $\psi_0^+$ and the number ($n_-$) of fermionic zero modes $\psi_0^-$ [1]. It is often a useful indicator of SUSY breaking: if $\Delta \neq 0$ then there must be at least one zero mode and so SUSY is unbroken; if $\Delta = 0$ then more information is needed, as this could be either because there are no bosonic or fermionic zero modes at all (in which case SUSY is broken), or it could be because $n_+ = n_- \neq 0$ (in which case SUSY is unbroken). Much of the power of the index method comes from the fact that the index may be reliably and easily computed in many cases, both in SUSY quantum mechanics and in SUSY field theory. To a large extent $\Delta$ is independent of the parameters (masses, couplings, volume, ...) of the theory; a property that allows the computation of $\Delta$ in special parametric regimes. Several direct applications of this are discussed in [1].

This insensitivity to fine-tuning of parameters must, of course, be qualified. It is assumed that under the parametric changes the energy eigenstates move about continuously in energy — they do not suddenly appear in, or disappear from, the Hilbert space. The simplest example is to take the superpotential to be $W(x) = x - a$, where $a$ is some constant. This is just the harmonic oscillator system and it is clear that $H_-$ has a normalizable zero mode for any finite value of the parameter $a$, while $H_+$ has no normalizable zero mode for any value of $a$. So SUSY is unbroken for all values of the parameter $a$.

On the other hand, with $W(x) = x^2 - a$, it is clear that neither $H_+$ nor $H_-$ has a normalizable zero mode for any value of $a$ and so SUSY is broken for all $a$. With $a > 0$ this is an example of dynamical SUSY breaking because the tree-level potential $W^2(x)$ does have zeros [1].

These simple examples with discrete spectrum may be generalized to include also continuum states. Consider

$$W(x) = \tanh x - a$$  \hspace{1cm} (5)

For $-1 < a < 1$, the fermionic hamiltonian $H_-$ has a normalizable zero mode $\psi_0^-(x) = e^{-a/2} \sech x$, while $H_+$ has no zero mode. Thus SUSY is unbroken. But for $|a| \geq 1$ the zero mode $\psi_0^-(x)$ of $H_-$ becomes non-normalizable; there are therefore no zero modes, and SUSY is therefore broken. This discontinuous change comes about because for $|a| \geq 1$ the asymptotic limits of the superpotential have changed from having opposite signs to having the same sign. Correspondingly the behavior of $V_-$ at $|x| = \infty$ has been altered. Nevertheless, in this example there is still a finite range $(-1, 1)$ of $a$ for which SUSY remains unbroken. On the other hand, consider the superpotential

$$W(x) = \tanh^2 x - a$$  \hspace{1cm} (6)

Clearly there is no $a$ for which either $H_-$ has a zero mode, and so SUSY is broken for all values of the parameter $a$. We shall return to this example later.

For periodic potentials the situation is rather different. The criterion for SUSY breaking reduces, as before, to the question of whether the zero modes in (4) are elements of the Hilbert space. For nonperiodic systems on the real line this is a question of normalizability of these zero-mode wavefunctions. This can be phrased in terms of the asymptotic limits of $W(x)$, or in terms of the number of zeros of $W(x)$ being odd or even, or in terms of whether
$W(x)$ is an odd or even function [1]. For periodic systems, normalizability is not the issue; rather, the zero-modes in (4) must be Bloch functions, and since they are zero energy they must in fact be periodic functions (with the same period as the potential) [4]. It is easy to see from (4) that this translates into the requirement [3] that the superpotential $W(x)$ satisfy:

$$\int_{\text{period}} W(x) = 0 \quad (7)$$

This condition has two simple, but significant, consequences. First, suppose the superpotential $W$ is such that $\int_{\text{period}} W \neq 0$. Then we can always arrange for $\int_{\text{period}} W = 0$ simply by subtracting an appropriate finite constant from $W$. Conversely, suppose the condition (7) is satisfied; then we have no freedom to shift $W$ by any finite constant $a$, without breaking SUSY.

We now illustrate these consequences with some examples that generalize those already mentioned for the nonperiodic case. First, consider the elliptic superpotential (see [3])

$$W(x) = m \text{sn}(x|m) \text{cd}(x|m) - a \quad (8)$$

where $\text{sn}(x|m)$ and $\text{cd}(x|m) = \text{cn}(x|m)/\text{dn}(x|m)$ are standard Jacobi elliptic functions [5], and $m$ is the elliptic parameter ($0 < m \leq 1$). The superpotential (8) has period $2K(m)$, where $K(m)$ is the "real elliptic quarter period". Note that when $m = 1$ the Jacobi functions reduce to hyperbolic functions: $\text{sn}(x|1) = \tanh x$ and $\text{cd}(x|1) = 1$. Thus, the superpotential in (8) reduces to $W = \tanh x - a$ which is just the example discussed earlier in (5).

Now, since $\frac{d}{dx}\left[ \log \text{dn}(x|m) \right] = -m \text{sn}(x|m) \text{cd}(x|m)$, and $\text{dn}(x|m)$ has period $2K(m)$, we find that

$$\int_{\text{period}} W = -2mK(m) \quad (9)$$

This only vanishes for $a = 0$, in which case SUSY is unbroken [3]. SUSY is broken for any nonzero value of the parameter $a$. Thus the periodic model in (8) is more sensitive to fine-tuning of the parameter $a$ than is the nonperiodic model with superpotential (5). This occurs even though there is no noticeable effect in the periodic system on the number of zeros of $W$, or on the values of $W(x)$ at the edges of a period, when $a$ deviates from 0.

Next, consider the periodic superpotential, with period $K(m)$,

$$W(x) = m^2 \text{sn}^2(x|m) \text{cd}^2(x|m) - a \quad (10)$$

When $m = 1$, $W$ reduces to $\tanh^2 x - a$, which coincides with the earlier example (6) for which SUSY was always broken, for all values of the parameter $a$. Here the situation is different – to determine whether SUSY is broken or not we look to the condition (7).

Note the following facts (see [5]): (i) $\int W = (2 - m - a)x + m \text{sn}(x|m) \text{cd}(x|m) - 2E(x|m)$, where $E(x|m)$ is the elliptic integral of the second kind: $E(x|m) = \int_0^x \text{dn}(s|m) ds$; (ii) under a period shift, $E(x + K(m)m) = E(x|m) - m \text{sn}(x|m) \text{cd}(x|m) + E(m)$, where $E(m)$ is the complete elliptic integral of the second kind: $E(m) = E(K(m)m)$.

Using these two facts, we find that

$$\int_{\text{period}} W = (2 - m - a)K(m) - 2E(m) \quad (11)$$

Thus, choosing

$$a = 2 - m - 2E(m)/K(m) \quad (12)$$

leads to unbroken SUSY. With this choice for the parameter $a$, both $H_\pm$ have periodic Bloch zero-modes,

$$\psi_0(\pm) = \exp\left\{ \pm \left[ m\text{sn}(x|m)\text{cd}(x|m) - 2Z(x|m) \right] \right\} \quad (13)$$

where $Z(x|m)$ is the Jacobi zeta function, which is related to the elliptic integral $E(x|m)$ through $E(x|m) = Z(x|m) + xE(m)/K(m)$.

Fig. 1 contains plots of both $W$ and $W^2$, showing that the tree-level potential $W^2$ has two zeros within each period. These plots have been made for the particular choice (12) for the parameter $a$. However, these plots are representative. In fact, if $0 < a < m^2/(1 + \sqrt{1 - m^2})$, $W^2$ always has two zeros within each period. The choice (12) falls within this range. The corresponding bosonic and fermionic potentials (1) are plotted in Fig. 2. Notice that the two potentials are simply parity reflections of one another – they are "self-isospectral" in the sense of [3]. The zero-mode wavefunctions (13) are plotted in Fig. 3. Notice that they are smooth, bounded, periodic, and have no zeros.

This model has an interesting $m \to 1$ limit. (Note that the period $K(m)$ diverges logarithmically [5] as
Fig. 1. The superpotential $W(x)$ (thin line) in (10), and its square $W^2(x)$ (thick line). These plots are for the elliptic parameter $m = .999$, for which the period is $K(m) = .84$. Notice that $W^2$, which is the tree-level potential, has two zeros within each period.

$m \to 1$.) With the parameter choice in (12), SUSY is unbroken for any $m < 1$. However, when $m = 1$, $W = -\text{sech}^2 x$, which is an even function on the real line, so that SUSY is broken. Indeed, the zero modes in (13) become $\psi_0^\pm = \exp(\mp \tanh x)$, neither of which is normalizable on the real line. This type of discontinuous behavior is interesting because we could alternatively (as was done for the model (8) in [6]) consider this not as a periodic potential model, but as a model defined on a finite interval with periodic boundary conditions, which would therefore have a discrete (rather than a band) spectrum. This type of regularization in a finite spatial volume is a common computational device. The conventional wisdom [1] is that if SUSY is shown to be unbroken in any finite volume, then this will persist in the infinite volume limit. Here, however, SUSY becomes broken in the infinite volume limit (i.e. when $m = 1$), even though it is unbroken for any finite volume (i.e. for any $m < 1$). How can this be happening? The answer is that the zeros of the tree-level potential $W^2$ within a single period, $-\frac{K(m)}{2} < x \leq \frac{K(m)}{2}$, (see Fig. 1) disappear to infinity in the infinite volume limit (i.e. as $m \to 1$). Indeed, when $m = 1$, $W^2 = \text{sech}^4 x$, which has no zeros at all. Thus, these vacua are receding to infinity and no longer play any role in the Hilbert space of the infinite volume theory. It is interesting that this type of behavior, with vacua disappearing to infinity, appears in the massless limit of SUSY QCD [7].

To conclude, we have shown that SUSY quantum mechanics models with periodic superpotentials are more sensitive to tuning of the parameters than are the more familiar nonperiodic models. Since the generic insensitivity of the Witten index to fine-tuning of parameters is often invoked in investigations of SUSY breaking in field theories, it would be interesting to learn whether these periodic models have field theoretic analogues.

**Acknowledgements**

This work has been supported by the DOE grant DE-FG02-92ER40716.00 (G.D.), and by the NSF (J.M.) through the R.E.U. (Research Experience for Undergraduates) Grant PHY-9424125 (W. Stwalley, P.I.) at the University of Connecticut.
References

Direct messenger-matter interactions in gauge-mediated supersymmetry breaking models

Tao Han, Ren-Jie Zhang

Department of Physics, University of California, Davis, CA 95616, USA
and Department of Physics, 1150 University Avenue, University of Wisconsin, Madison, WI 53706, USA

Received 26 February 1998
Editor: M. Dine

Abstract

We categorize general messenger and matter interactions in gauge-mediated SUSY breaking models by an $R$-parity for the messengers and study their phenomenological consequences. The new interactions may induce baryon- and lepton-number violating processes as well as flavor-changing neutral currents. Bounds on the couplings from low-energy data are generally weak due to the large messenger mass suppression, except for the constraint from proton decay. The soft masses for the scalar particles receive negative corrections from the new interactions. Consequently, in certain region of SUSY parameter space the $\mu$-parameter is greatly reduced. The pattern of radiative electroweak symmetry breaking, SUSY particle mass spectrum and decay channels are also affected, leading to observable experimental signature at the current and future colliders. © 1998 Elsevier Science B.V. All rights reserved.

PACS: 12.60.Jv; 12.60.-i; 14.80.-j; 14.80.Ly

1. Introduction

A model with gauge-mediated supersymmetry breaking (GMSB) [1] is a simple and well-motivated version of the minimal supersymmetric extension of the Standard Model (MSSM). In addition to the observable sector and a supersymmetry (SUSY) breaking hidden sector, the model also possesses messenger fields which mediate the SUSY breaking to the observable fields via the SM gauge interactions. The “minimal” model has a pair of messengers transforming under the SU(5) representation $\mathbf{5} + \bar{\mathbf{5}}$, decomposed as color triplets $(D + \bar{D})$ and weak doublets $(L + \bar{L})$. They couple to a gauge singlet field $S$ through a superpotential

$$W_{\text{minimal}} = \lambda (SD\bar{D} + SLL),$$

where $S$ acquires non-zero vacuum expectation values for both its scalar component $\langle S \rangle$ and auxiliary component $F$.

There are several attractive features in this minimal model. First, all supersymmetric particle masses are determined by two parameters: the messenger scale $M = \lambda \langle S \rangle$ (the messenger fermion mass) and the effective SUSY breaking scale $A = F/\langle S \rangle$. The
gaugino and scalar soft masses are given, at one- and two-loop level respectively, by [1]

\[ M_i(M) = \frac{\alpha_i(M)}{4\pi} \Lambda, \quad i = 1,2,3, \quad (2) \]

\[ \tilde{m}^2(M) = 2 \sum_{i=1}^{3} C_i \left( \frac{\alpha_i(M)}{4\pi} \right)^2 \Lambda^2, \quad (3) \]

where \( C_i \)'s are 4/3, 3/4 for the fundamental representations of SU(3), SU(2) and 3Y^2/5 for U(1). If \( \Lambda \sim \mathcal{O}(10-100 \text{ TeV}) \), the SUSY particles (sparticles) can have a desirable mass spectrum of \( \mathcal{O}(100 \text{ GeV}) \).

Second, since the scalar masses are degenerate in the family space, the flavor-changing neutral current (FCNC) and CP-violation in SUSY sector are generally small. Finally, the gravitino mass is typically of \( \mathcal{O}(10^{-100} \text{ TeV}) \), the SUSY particles sparticles can have a desirable mass spectrum of \( \mathcal{O}(100 \text{ GeV}) \).

However, there may be cosmological problems in the messenger sector [3]. Due to the conserved “messenger number” in Eq. (1), the lightest messenger particle (LMP) is stable. Although naturally neutral [4] in most of the SUSY parameter space, the LMP would have to be lighter than a few TeV in order not to overclose the Universe [3,4] in the standard inflationary cosmology. On the other hand, direct searches have already excluded a scalar dark matter particle with a mass less than about 3 TeV at a 90% confidence level, assuming it accounts for more than about 30% of a galactic halo with local density 0.3 GeV/cm^3 [5]. One would have to introduce the messenger-Higgs mixing, along with a gauge singlet, to evade the direct dark matter detection [4]. However, with those additional interactions, one may run into the \( \mu-B_\mu \) problem [1,6,7]. Besides, such a low-mass LMP needs a certain degree of fine-tuning.

One possible solution to the problem is to abandon the messenger number conservation by introducing direct messenger-matter interactions. In fact, it is natural to consider this possibility since the messengers intrinsically carry the SM gauge quantum numbers. In this paper, we study general interactions between the messengers and MSSM fields. In Section 2 we present the most general superpotential in the framework of the minimal GMSB model and categorize different terms with an \( R \)-parity for the messengers. We then derive the effective Lagrangian at low energies and examine the current experimental constraints on the couplings in Section 3. We also study the theoretical implications on the SUSY particle mass spectrum, the electroweak symmetry breaking (EWSB) and the \( \mu \)-parameter. In Section 4 we make some general remarks and draw our conclusion.

2. General messenger-matter interactions

The most general superpotential with direct messenger-matter interactions allowed by the SM gauge symmetry is

\[ W_{\text{mix}} = H_D L_D \tilde{E} + H_D Q D_4 + LL \tilde{E} + QL \tilde{T} D_4 + QL \tilde{U} U \\
+ E \tilde{T} D_4 + LQ \tilde{D}_4 + QD_4 + \tilde{U} \tilde{T} \tilde{D}_4 \\
+ H/ L_4 + H_D \tilde{L}_4 + \tilde{L}_4 + \tilde{D}_4 + QL \tilde{D}_4, \quad (4) \]

where we have suppressed the Yukawa coupling constant for each term and the generation indices for the superfields. A subscript “4” has been introduced for the messenger fields in Eq. (1) as \( (D_4, L_4) \) and \( (\tilde{D}_4, \tilde{L}_4) \), in analogue to the three-generation matter fields.

At this superpotential level, the bilinear terms can be rotated away at no cost by properly redefining the superfields, so we will not consider them any further. We will also ignore the state mixing among the messengers and MSSM fields associated with those rotations. The last term in Eq. (4) is the only one which involves two messenger fields. When examining its physical consequences at low energies by integrating out the heavy messenger fields, the resulting operators would be more suppressed. Although this term respects the SM gauge symmetry, it does not naturally arise in an SU(5) unification theory. We will not discuss this term further.

To classify the remaining terms in Eq. (4), we recall that the superfields \( (H_D, H_D) \) have a positive
(matter) R-parity assignment while the others $(L, E, Q, U, D)$ are negative. It follows that the first two terms in Eq. (4) have different R-parity property from the rest. It is therefore convenient to categorize these two groups by their R-parities. If we formally require the R-parity conservation, we then can generalize the ordinary R-parity to the messengers. The two possibilities are: messenger superfields with a positive R-parity, which we will call the $M^+$-model, and messenger superfields with a negative R-parity, the $M^-$-model.

2.1. The $M^+$-model

If we assign the messenger superfields $(D_{ij}^a, L_j^a)$ and $(\overline{D}_i^a, \overline{L}_i^a)$ with a positive R-parity, then R-parity invariance leads to the following interaction terms:

$$W_{\text{mix}}^+ = y_{ij} E_i L_j L_4 + y_{ij}' \overline{D}_i Q_j L_4$$

$$+ y_{ij}'' \overline{D}_i \overline{E}_j Q \overline{L}_4 + \lambda_{ij}^r E_i \overline{D}_j D_4 + \lambda_{ij}' L_j Q \overline{D}_4$$

$$+ \frac{1}{2} \lambda_{ij}'' Q_j Q \overline{D}_4 + \lambda_{ij}'' \overline{D}_j \overline{D}_4,$$  

(5)

where $y_{ij}, y_{ij}', y_{ij}''$ and $\lambda_{ij}^r, \lambda_{ij}', \lambda_{ij}''$ are Yukawa couplings naturally of order one $^2$, with $i, j = 1, 2, 3$ as generation indices. Note that $\lambda_{ij}^r = - \lambda_{ij}'$. According to the conventional R-parity assignment for component fields, $(-1)^{S+3(B-L)}$, where $S$ is the particle spin, $B$ and $L$ the baryon- and lepton-number respectively, the R-parity so assigned corresponds to that assuming zero $B$- and $L$-numbers for the messenger superfields. The first three terms conserve $B$ and $L$, but generate FCNC processes in general. Although Eq. (5) preserves R-parity by definition, the $\lambda^r, \lambda'$ terms in (5) violate $B$ and $L$ by $\Delta B = 1/3$ and $\Delta L = 1$, and the $\lambda^r, \lambda'$ terms violate $B$ by $\Delta B = 2/3$. Simultaneous existence of $\lambda^r$ and $\lambda^r, \lambda'$ may induce abrupt proton decay. $^3$ We will discuss this point in the next section. Note that the terms of $y_{ij}, \lambda_{ij}^r, \lambda_{ij}''$ are the direct analogue to those R-parity violating interactions of $\lambda_{ij}, \lambda_{ij}', \lambda_{ij}''$ in the MSSM.

2.2. The $M^-$-model

If we instead assign the messenger superfields with a negative parity, then R-parity invariance excludes the terms in Eq. (5) and we are left with the two terms involving Higgs superfields

$$W_{\text{mix}}^- = y_i H_D L_i E_i + y_i' H_D Q_i \overline{D}_4.$$  

(6)

This R-parity assignment is equivalent to assuming the messenger superfields $(\overline{D}_i^a, L_j^a)$ to carry the same $B$- and $L$-numbers as the MSSM superfields $(\overline{D}_i, L_j)$. Although these two terms do not induce any $B$- or $L$-violating processes, the messenger couplings to the MSSM fields will mediate FCNC processes $[7,9]$ if more than one $y_i$ or $y_i'$ coexists.

It is interesting to note that in either R-parity assignment ($M^+$ or $M^-$), R-parity invariance forbids the $QL \overline{D}_4$ term in Eq. (4).

3. Physical implications

3.1. Effective lagrangian and low-energy constraints

By integrating out heavy messengers, we can obtain low-energy effective Lagrangian in terms of $1/M^2$ expansion. For simplicity, we only examine the leading operators by assuming $\Lambda/M \ll 1$ and the sparticles and Higgs bosons to be much heavier than the energy scale considered.

In the $M^+$-model of Eq. (5), the first three terms ($y$-couplings) result in the following four-fermion operators:

$$\mathcal{L}_{\text{lag}}^+ = \frac{y_{ij} y_{ij}'}{2 M^2} \left( e_{ij} e_{ij}^* \overline{e}_{ij} \gamma_\mu e_{iL} \gamma_\mu e_{jL} \right.$$  

$$+ e_{ij} e_{ij}^* \overline{e}_{ij} \gamma_\mu e_{iL} \gamma_\mu \nu_{iL} \right)$$

$$+ \frac{y_{ij} y_{ij}''}{2 M^2} \left( \overline{d}_{ij} \gamma_\mu d_{ij}^* \overline{d}_{ij} \gamma_\mu \nu_{iL} \right)$$

$$+ \frac{y_{ij} y_{ij}'}{2 M^2} \left( \overline{u}_{ij} \gamma_\mu u_{ij}^* \overline{u}_{ij} \gamma_\mu \nu_{iL} \right)$$

$$+ \frac{1}{2} \frac{y_{ij} y_{ij}''}{2 M^2} \left( \overline{e}_{ij} \gamma_\mu e_{ij}^* \overline{e}_{ij} \gamma_\mu \nu_{iL} \right).$$
suppression by the large messenger mass can be naturally of order one. However, if we consider the effective four-fermion operators for the last four terms (λ-couplings) in Eq. (5) to leading order of $1/M^2$,

$$\mathcal{L}_\lambda^\ast = -\frac{2}{M^2} \left( \lambda^\ast_{ij} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

\right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

\right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

\right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

\right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

\right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

where $\alpha, \beta$ are color indices. Similarly, we can obtain the effective four-fermion operators for the last four terms (λ-couplings) in Eq. (5) to leading order of $1/M^2$,

$$\mathcal{L}_\lambda^\ast = -\frac{2}{M^2} \left( \lambda^\ast_{ij} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

\right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

\right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

\right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

\right) + \frac{\lambda^\ast_{ij} \lambda^\ast_{k\ell} \gamma^\mu e_{iR} \gamma_{\nu_{iL}} u_{jL} + \text{h.c.} \right) \right)

These effective operators are similar to those obtained from the $R$-parity violating interactions in the MSSM [10] and may lead to very rich physics. The coupling coefficients ($\gamma'$s and $\lambda'$s) are considered to be naturally of order one. However, if we consider only one term at a time with given generation indices in Eq. (5), there is essentially no significant experimental constraint on them because of the suppression by the large messenger mass $M \approx \mathcal{O}(100$ TeV). For the same reason, none of the terms would lead to observable signature in the current and near-future experiments. Even when several terms with different generation indices coexist, most effects of those operators in Eqs. (7) and (8) are still rather weak in general. For example, the modifications on charged current universality and on various $\tau$, $D$ and $B$ decays are too small to be observable unless the couplings $\gamma_{ij}, \lambda_{ij}^\ast, \lambda_{ij} > \mathcal{O}(100)$. On the other hand, there are processes from rare and SM forbidden decays and from neutral meson mixing that can be sensitive to test certain operators with the accuracy of current and near-future experiments. For instance, considering $\mu \rightarrow e\gamma$ with one-loop diagrams via the virtual messenger exchange, we obtain

$$|\gamma_{ij}, \lambda_{ij}^\ast, \lambda_{ij} > \mathcal{O}(100)$$

$$|\gamma_{ij}, \lambda_{ij}^\ast, \lambda_{ij} > \mathcal{O}(100)$$

The most stringent bound on $y'$ comes from the $K_L-K_S$ mass difference:

$$|\gamma_{ij}, \lambda_{ij}^\ast, \lambda_{ij} > \mathcal{O}(100)$$

which is at a very interesting level to constrain the theory. The $D^0$ mass difference constrains the operators $\gamma_{ij}, \lambda_{ij}^\ast, \lambda_{ij} > \mathcal{O}(100)$, while the $B^0$ mass difference bounds the operators with the third generation index $\gamma_{ij}$. It is also interesting to note that the operators $\gamma_{ij}, \lambda_{ij}^\ast, \lambda_{ij} > \mathcal{O}(100)$ contributing to $K_L \rightarrow \ell^K, \ell^K$ are of the nature of a pseudoscalar current, so there is no explicit lepton mass dependence in the decay width, unlike the $R$-parity violating interactions [11] where there is essentially no sensitivity to new physics for the decay $\bar{K}_L \rightarrow e^+e^-$. This feature may serve as a criterion to distinguish different models if a signal beyond the SM is observed.

The last three operators in Eq. (8) mediate proton decay such as $p \rightarrow e^+\pi^0(K^0), \mu^+\pi^0(K^0)$ and $\nu\pi^+(K^+)$. Requiring the proton lifetime to be larger than $10^{32}$ years puts very stringent bounds to them: products of two appropriate couplings are restricted at order of $10^{-21}$ for a 100 TeV messenger. It is therefore unlikely for the operators $\lambda', \lambda''$ to coexist. In Table 1, we summarize the meaningful bounds on the products of two different couplings in Eqs. (7) and (8), along with the corresponding exper-
immental data [12]. Future high precision measurements would explore the operators to a more significant level.

In the $M^*$-model, by integrating out the messenger and Higgs fields, one may obtain fermionic bilinear terms. However, they are not only suppressed by the heavy masses $M$ and $m_h$, but also by chirality. They are generally small and we will not discuss them.

### 3.2. Sparticle mass spectrum

While the masses of MSSM fermions are protected from the radiative corrections either by a chiral symmetry or by a continuous $R$-symmetry, the scalar masses squared receive large negative corrections from the messenger-matter interactions in Eqs. (5) and (6). They are given by, for the $M^*$-model,

$$
\delta m^2_{\ell_{ij}} = \frac{1}{16\pi^2} \sum_{k=1}^3 y_{ik} y_{jk}^* M^2 u(\Lambda/M),
$$

(11)

$$
\delta m^2_{k_{i}} = \frac{1}{8\pi^2} \sum_{k=1}^3 y_{ik} y_{jk}^* M^2 u(\Lambda/M),
$$

(12)

$$
\delta m^2_{j_{i}} = \frac{1}{16\pi^2} \sum_{k=1}^3 (y_{ik} y_{jk}^* + y_{ik}^* y_{jk})

\times M^2 u(\Lambda/M),
$$

(13)

$$
\delta m^2_{k_{i}} = \frac{1}{8\pi^2} \sum_{k=1}^3 y_{ik} y_{jk}^* M^2 u(\Lambda/M),
$$

(14)

$$
\delta m^2_{\ell_{i}} = \frac{1}{8\pi^2} \sum_{k=1}^3 y_{ik} y_{jk}^* M^2 u(\Lambda/M),
$$

(15)

where we have ignored the $\lambda$-couplings in Eq. (5), and for the $M^*$-model [7,9],

$$
\delta m^2_{\ell_{i}} = \frac{1}{8\pi^2} y_{i} y_{i}^* M^2 u(\Lambda/M),
$$

(16)

$$
\delta m^2_{j_{i}} = \frac{1}{16\pi^2} y_{i} y_{i}^* M^2 u(\Lambda/M),
$$

(17)

$$
\delta m^2_{\ell_{i}} = \frac{1}{16\pi^2} \sum_{i=1}^3 (|y|^2 + 3|\nu|^2) M^2 u(\Lambda/M),
$$

(18)

where the function

$$
\tan(x) = \ln(1 + x^2) + \frac{x}{2} \ln \frac{1 + x}{1 - x}.
$$

For small $x$, the function has an expansion $-x^4/6$, and it monotonically decreases as $x$ goes to 1.

For $\Lambda/M \ll 1$ ($x \to 0$), the corrections are suppressed by $(\Lambda/M)^2$, so that there is no significant difference from the minimal model for sparticle spectrum. On the other hand, for $\Lambda/M \approx 1$ ($x \to 1$), the function $u$ is very negative and the corrections to the mass can be very substantial. Significant upper limits on the Yukawa couplings may be obtained by requiring that these negative corrections do not change the sign of the scalar mass squared. In GMSB models, the slepton and Higgs soft masses are gener-
ically smaller than the squark soft masses, so the
tightest bounds come from Eqs. (11), (12), (16) and
(18). As an illustration, we choose
\[ \Lambda = 100 \text{ TeV}, \tan \beta = 2 \quad \text{and} \quad \mu > 0. \quad (20) \]
By requiring the scalar mass squared to remainpositive, we find that typical upper bounds on \( \Sigma_{i=1}^{3} |y_i|^2 \) in the \( M^- \)-model for several \( M/\Lambda \) values are
\[ M/\Lambda = \begin{array}{cccc} 1.25 & 5 & 10 & 30 \\
\sum_{i=1}^{3} |y_i|^2 & < & 10^{-3} & 0.03 & 0.15 & 1.2 . \end{array} \quad (21) \]
Similar constraints are also obtained for \( \Sigma_{i=1}^{3} |y_i|^2 \) and for the couplings in the \( M^- \)-model. Although the bounds obtained here depend on the model parameter \( M/\Lambda \), they are the only upper bounds available on individual couplings. They are therefore complementarily to those extracted from the low-energy data in the previous section.

The negative corrections to the scalar masses squared can also induce misalignments of the fermion-sfermion mass matrices and as a result, the flavor-changing neutral currents. A study [9] found that bounds on the products of two couplings from \( \mu \to e\gamma \) and \( \mu-e \) conversion can be as strong as \( 10^{-5} \) for \( A/M \approx 1 \). However, the bounds obtained there depend again sensitively on the parameter choice, and they are much looser for \( A/M \ll 1 \).

### 3.3. Electroweak symmetry breaking and the \( \mu \)-parameter

One of the most important features in SUSY theories is the radiative generation of the electroweak symmetry breaking (EWSB) [14]. At the scale \( M_{\text{SUSY}} \) where the EWSB is imposed, the Higgs soft mass squared \( m_{h_u}^2 \) is approximately given by the solution to the one-loop renormalization group equation
\[ m_{h_u}^2(M_{\text{SUSY}}) = m_{h_u}^2(M) - \frac{3 \lambda_\mu^2}{8 \pi^2} \left( m_{\tilde{q}_3}^2 + m_{\tilde{\ell}_3}^2 \right) \ln \left( \frac{M}{M_{\text{SUSY}}} \right), \]
where for simplicity \( \lambda_\mu \) term has been neglected. The large top-quark Yukawa coupling \( \lambda_t \) can drag \( m_{h_u}^2 \) at \( M_{\text{SUSY}} \) negative, thus triggers the EWSB.

However, the messenger-matter interactions give large negative corrections to the scalar masses squared, as seen in Eqs. (11)–(15) and Eqs. (16)–(18). When the third generation squark masses become small, the Higgs mass squared is less negative, and the EWSB can be changed significantly or may not even occur. A more restrictive version of the EWSB condition in a supersymmetric theory is usually expressed in the following (tree-level) equation which also determines the \( \mu \)-parameter
\[ \mu^2 = \frac{m_{h_u}^2 - \tan^2 \beta m_{\tilde{h}_u}^2}{\tan \beta - 1} = \frac{M^2}{2}. \]

Requiring the model to yield a desirable pattern of EWSB would put additional constraints on the couplings. We consider the mass correction effects on EWSB in \( M^- \)-model to Eqs. (22) and (23). We choose the SUSY parameters as in Eq. (20). To maximize the effects from messenger-matter interactions, we also choose that \( M/\Lambda = 1.25 \) and \( y_{33} = y_{23} \). We have run the coupled two-loop renormalization group equations [15] of soft masses to the scale \( M_{\text{SUSY}} = m_{\tilde{q}_3}^{\text{GM}} + m_{\tilde{q}_3}^{\text{mix}}, \) where \( m_{\tilde{q}_3}^{\text{GM}} \) and \( m_{\tilde{q}_3}^{\text{mix}} \) repre-

![Fig. 1. Representative masses as functions of the messenger-matter coupling \( y_{33}^2 \). The hatched region does not have the right EWSB. Shown in the plots are the third generation squarks \( \tilde{q}_3 \), the \( \mu \)-parameter, the lightest neutralino \( \tilde{\chi}_1^0 \) and the lightest CP-even and CP-odd Higgs bosons \( h, A \).](image-url)
sent the contributions to the third generation squark masses from the minimal GMSB model (Eq. (3)) and messenger-matter interactions (Eqs. (13)–(15)) respectively. At $M_{\text{SUSY}}$ we impose the EWSB condition and calculate all physical masses and the $\mu$-parameter consistently to the full one-loop order. In Fig. 1 we show our results for the SUSY particle mass spectrum. The hatched region is where the EWSB does not occur. This can be anticipated from looking at the decreasing squark masses, which eventually become too small to drive the Higgs mass squared of Eq. (22) negative. The limits on the individual Yukawa couplings obtained here are similar to that in Eq. (21) and are complementary to those from the low-energy experiments discussed in the previous section. We note that the $\mu$-parameter decreases as the coupling increases.

Models with direct messenger-matter interactions can display mass spectrum with very different characteristics from that of the minimal GMSB model.

Table 2
Representative masses for the two models: minimal GMSB with $y'_{33} = y'_{13} = 0$, and that for $y'_{33} = y'_{13} = 0.032$. Also shown in the Table are the $\mu$-parameter, the three gaugino soft masses $M_{1,2,3}$, branching ratios for $\tilde{\chi}_1^0 \rightarrow \gamma \tilde{G}$ and $\tilde{\chi}_1^0 \rightarrow h \tilde{G}$, and $\tilde{\chi}_2^0$ decay length.

<table>
<thead>
<tr>
<th>(GeV)</th>
<th>$y'<em>{33} = y'</em>{13} = 0$</th>
<th>$y'<em>{33} = y'</em>{13} = 0.032$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>619</td>
<td>152</td>
</tr>
<tr>
<td>$M_1$</td>
<td>159</td>
<td>157</td>
</tr>
<tr>
<td>$M_2$</td>
<td>299</td>
<td>299</td>
</tr>
<tr>
<td>$M_3$</td>
<td>820</td>
<td>846</td>
</tr>
<tr>
<td>$m_{\tilde{\chi}_1^0}$</td>
<td>158</td>
<td>143</td>
</tr>
<tr>
<td>$m_{\tilde{\chi}_2^0}$</td>
<td>316</td>
<td>169</td>
</tr>
<tr>
<td>$m_{\tilde{\chi}_3^0}$</td>
<td>625</td>
<td>172</td>
</tr>
<tr>
<td>$m_{\tilde{\chi}_4^0}$</td>
<td>631</td>
<td>325</td>
</tr>
<tr>
<td>$m_{\tilde{\chi}_5^0}$</td>
<td>908</td>
<td>888</td>
</tr>
<tr>
<td>$m_{\tilde{\chi}_6^0}$</td>
<td>158</td>
<td>163</td>
</tr>
<tr>
<td>$m_{\tilde{\chi}_7^0}$</td>
<td>631</td>
<td>325</td>
</tr>
<tr>
<td>$m_{\tilde{\chi}_8^0}$</td>
<td>94</td>
<td>83</td>
</tr>
<tr>
<td>$m_{\tilde{\tau}_1}$</td>
<td>802</td>
<td>429</td>
</tr>
<tr>
<td>$m_{\tilde{\tau}_2}$</td>
<td>1099</td>
<td>627</td>
</tr>
<tr>
<td>$m_{\tilde{\tau}_3}$</td>
<td>1000</td>
<td>495</td>
</tr>
<tr>
<td>$m_{\tilde{\tau}_4}$</td>
<td>1090</td>
<td>606</td>
</tr>
<tr>
<td>$m_{\tilde{\tau}_5}$</td>
<td>1083</td>
<td>531</td>
</tr>
<tr>
<td>BR($\tilde{\chi}_1^0 \rightarrow \gamma \tilde{G}$)</td>
<td>0.90</td>
<td>0.65</td>
</tr>
<tr>
<td>BR($\tilde{\chi}_1^0 \rightarrow h \tilde{G}$)</td>
<td>$\sim 10^{-5}$</td>
<td>0.26</td>
</tr>
<tr>
<td>$c\tau (\mu m)$</td>
<td>84</td>
<td>60</td>
</tr>
</tbody>
</table>

To demonstrate this point, we show a representative mass spectrum in Table 2, where we choose $y'_{33} = y'_{13} = 0.032$ and all other parameters are taken to be the same as Fig. 1. For comparison, we also show that for the minimal GMSB model with no direct messenger-matter interactions. The masses for the third generation squarks, the neutralino/chargino and especially the Higgs bosons are all significantly lighter than those for the minimal model. The lightest neutralino $\tilde{\chi}_1^0$ has a large Higgsino component, so it can decay to the light Higgs $h$ with a fairly large branching ratio of 26%. The decay length of $\tilde{\chi}_1^0$ becomes somewhat shorter as well. These interesting features may lead to very distinctive experimental signature in the current and future collider experiments. The slepton masses can be decreased significantly as well depending on the choice of the couplings $y'$s.

In the minimal GMSB model, because the squarks are much heavier than other sparticles, one finds that $|m_{\tilde{H}_u}|$ and therefore $\mu$ are typically much larger than $M_\mu$. This renders a difficult balance of Eq. (23) and is usually referred to as the fine-tuning problem in GMSB models [16]. In the presence of messenger-matter interactions, one should expect that $\mu$ is generally smaller (as seen from Fig. 1), and the fine-tuning problem should be less severe. We exam-
ine the dimensionless quantities as a measure of fine-tuning [17]:

\[ c(M_Z^2; \mu) = |\partial \ln M_Z^2 / \partial \ln \mu| \]

and

\[ c(M_Z^2; B_\mu) = |\partial \ln M_Z^2 / \partial \ln B_\mu| \]  

(24)

where \( B_\mu \) is the bilinear soft Higgs mass parameter. The results are shown in Fig. 2. Indeed the fine-tuning improves for a bigger coupling \( y_{33} \).

4. Discussions and conclusion

Before we draw our conclusions, several remarks are in order. First, in constructing the low-energy effective Lagrangian in the previous section, we have ignored terms proportional to \( \Lambda / M \). This is the case where the SUSY breaking effect in the messenger sector is much smaller than the messenger scale \( M \) itself. The terms proportional to \( \Lambda / M \) have essentially the same structure as those in Eqs. (7) and (8), with somewhat different combination of the Yukawa couplings. The physics implications are however very much similar.

Second, in principle, one can integrate out only the heavy messengers and obtain effective Lagrangians involving external sparticles. For example, terms in Eq. (5) could induce pair productions of \( l_i l'_j \) and \( q_i q'_j \) at lepton and hadron colliders, and those in Eq. (6) would give Higgs and Higgsino pair production. Although the strength of the new interactions is generically small, these distinctive processes may provide new experimental signature at future colliders.

Third, we have neglected the complication of the CKM matrix when deriving the couplings from Eq. (5) to Eqs. (7) and (8). It can be systematically included by performing the proper quark field rotation between the weak and mass eigenstates.

Finally, although we only concentrate on a pair of \( 5 + \bar{5} \) in this paper, the analysis can also be readily carried out for the cases of several \( 5 + \bar{5} \) pairs or a pair of \( 10 + \bar{10} \). In the case where the messengers are a pair of \( 10 + \bar{10} \), the LMP is charged and cannot be stable without causing problems in the standard inflationary cosmology. The direct messenger-matter interactions thus may necessarily occur. We should note that in these cases general mixing among messengers are also possible. Under certain assumptions, it is shown in Ref. [18] that hypercharge \( D \)-term contributions to the scalar particle masses can be generated at two-loop level, but these terms are generally much smaller than the one-loop contributions from the messenger and matter interactions.

In conclusion, we have constructed the direct messenger-matter interactions in the minimal GMSB model. The new interactions avoid the cosmological problem associated with the stable messenger particle, but they generally introduce \( B \) and \( L \) violating and FCNC processes. We obtain the low-energy effective Lagrangians by integrating out the heavy messenger fields as well as the sparticles. If we assume that the couplings are naturally of order one, we find that the constraints from the low-energy data are generally not very restrictive except for those leading to proton decay. On the other hand, certain combinations of the couplings may contribute to some flavor violating processes within the current and future experimental reach. We also show that the new interactions have negative contributions to the scalar particle masses. Consequently, one can generally reduce the value of the \( \mu \)-parameter and greatly change the pattern of EWSB. The fine-tuning problem associated with the \( \mu \)-parameter can be alleviated at most of the parameter space. The significantly lighter mass spectrum for the sparticles and Higgs bosons and the different sparticle decay pattern can result in distinctive experimental signature at the current and future colliders.

Acknowledgements

We would like to thank K. Cheung and R. Hempfling for helpful conversations and K. Matchev for comments on the manuscript. This work was supported by DOE grants DE-FG03-91ER40674 and DE-FG02-95ER40896.

References


Constraints on \( R \) parity and \( B \) violating couplings in gauge-mediated supersymmetry breaking models

Kiwoon Choi 1, Kyuwan Hwang 2, Jae Sik Lee 3

Department of Physics, Korea Advanced Institute of Science and Technology, Taejon 305-701, South Korea

Received 2 March 1998

Editor: M. Dine

Abstract

We consider the proton decay involving a light gravitino or axino in gauge-mediated supersymmetry breaking models to derive constraints on the \( R \) parity and baryon number violating Yukawa couplings. Bounds on all nine coupling constants are obtained by considering the decay amplitudes at one-loop order. © 1998 Elsevier Science B.V. All rights reserved.

PACS: 13.30.-a; 11.30.Fs; 12.60.Jv; 14.80.Mz

In supersymmetric models, there can be renormalizable gauge-invariant terms in the superpotential which violate the baryon number \( B \) or the lepton number \( L \). To avoid such terms, one usually introduces an additional discrete symmetry, the so-called \( R \) parity \((R_p = (-1)^{3B+L+y})\). Although \( R_p \) conservation leads to a consistent theory, there is no compelling theoretical reason to assume this symmetry. It is therefore an interesting possibility to have an explicit \( R \) violation which may lead to interesting phenomenological consequences [1]. In the minimal supersymmetric standard model, the most general \( R_p \)-violating superpotential is given by

\[
\frac{1}{2} \lambda_{ijk} L_i L_j E_{k}^c + \lambda'_{ijk} L_i Q_j D_{k}^c + \frac{1}{2} \lambda''_{ijk} U_i'^c D_j'^c D_k'^c, \quad (1)
\]

where \( L_i \) and \( Q_i \) are the \( SU(2) \)-doublet lepton and quark superfields and \( E_{i}^c, U_{i}'^c, D_{i}'^c \) are the singlet superfields, respectively. Here \( i,j,k \) are generation indices and we assume that possible bilinear terms \( \mu_i L_i H_u \) are rotated away. Obviously the first and second terms in (1) violate \( L \), while the third violates \( B \). Since \( \lambda_{ijk} = -\lambda_{jik} \) and \( \lambda'_{ijk} = -\lambda''_{ijk} \), \( R_p \) violations are described by 45 complex Yukawa couplings (9 in \( \lambda \), 27 in \( \lambda' \) and 9 in \( \lambda'' \)).

It is well known that the consideration of proton decay provides a very stringent constraint on the product of \( \lambda' \) and \( \lambda'' \):

\[ |\lambda'_{11k} \lambda''_{11k}|, |\lambda'_{22k} \lambda''_{22k}| \leq 10^{-24}, \quad (2) \]

where the squark masses are assumed to be around 1 TeV [2]. This bound has been obtained from a squark-mediated proton decay at tree level which does not involve heavy generation particles and thus applies for the particular combination of generation indices as is shown above. One may then expect that other products of \( \lambda' \) and \( \lambda'' \) are allowed to be large.
However it has been noted [3] that for any pair of $\lambda'$ and $\lambda''$, there is always at least one diagram relevant for the proton decay at one-loop level. This means that all products of $\lambda'$ and $\lambda''$ can be constrained by proton decay and a more detailed analysis leads to [3]

$$[\lambda' \cdot \lambda''] \leq 10^{-9},$$

for any pair of $\lambda'$ and $\lambda''$.

As was noted in [4,5], if there is a light fermion (lighter than the proton) which does not carry any lepton number, proton decay can be induced by a $B$ violating but $L$ conserving interaction alone, for instance by the $\lambda''$ couplings alone. Perhaps the most interesting class of models predicting such a light fermion are supersymmetric models in which supersymmetry (SUSY) breaking is mediated by gauge interactions [6]. In such models, the squark and/or gaugino masses, i.e. the soft masses in the supersymmetric standard model (SSM) sector, are given by $m_{\text{soft}} = (\alpha/\pi)^2 \Lambda_5$ where $\Lambda_5$ corresponds to the scale of spontaneous SUSY breaking and the model-dependent integer $n$ counts the number of loops involved in transmitting SUSY breaking to the supersymmetric standard model sector. On the other hands, the gravitino mass is suppressed by the Planck scale $M_p = 2 \times 10^{18}$ GeV, $m_{3/2} = \Lambda_5/M_p$. Assuming that $m_{\text{soft}}$ is at the weak scale and taking $n = 1 \sim 3$ for instance, we have $m_{3/2} = 10^{-1} \text{eV} \sim 10 \text{MeV}$ which is far below the proton mass.

Another interesting candidate for a light fermion without carrying any lepton number is the axino in supersymmetric models with a spontaneously broken global $U(1)_{PQ}$ symmetry [7]. If SUSY breaking is mediated by gauge interactions, the axino mass is given by $m_a = (\alpha/\pi)^n \Lambda_5^2 / F_p$ where $m$ is again a model-dependent (but typically not less than $n$) integer and $F_p$ denotes the scale of spontaneous $U(1)_{PQ}$ breaking [8]. Obviously in this case the axino can be lighter than the proton for a phenomenologically allowed $F_p \geq 10^{10}$ GeV. In other type of models in which SUSY breaking is transmitted by supergravity interactions, the gravitino mass is fixed to be of the weak scale order, however there is still a room for an axino lighter than the proton [9]. As was pointed out in Ref. [9], some supergravity-mediated models lead to $m_a \approx (m_{3/2}/M_p)^{1/2} = 1 \text{keV}$ for which the axino would be a good warm dark matter candidate [10].

In Ref. [5], proton decay involving a light gravitino or axino has been analysed at tree approximation to obtain a constraint on the $R_p$ and $B$ violating Yukawa coupling $\lambda''_{qL}$. Applying the naive dimensional analysis rule [11] for the hadronic matrix element of the effective 4-fermion operator induced by the tree diagram of Fig. 1, the following stringent bounds on $\lambda''_{112}$ (in the quark mass eigenstate basis) were obtained:

$$\lambda''_{112} \leq 5 \times 10^{-16} \left( \frac{\bar{m}}{300 \text{GeV}} \right)^2 \left( \frac{m_{3/2}}{1 \text{eV}} \right),$$

$$\lambda''_{112} \leq 7 \times 10^{-16} \left( \frac{\bar{m}}{300 \text{GeV}} \right)^2 \left( \frac{F_p}{10^{10} \text{GeV}} \right) \left( \frac{1}{c_q} \right),$$

where $\bar{m}$ denotes the squark mass which is presumed to be universal. Here the dimensionless coefficient $c_q$ describes the axino coupling to the light quarks $q = (u,d,s)$ and is of order one for Dine-Fischler-Srednicki-Zhitnitskii type axino, while it is of order $10^{-2} \sim 10^{-3}$ for hadronic-type axino. In this paper, we wish to extend the analysis of [5] by including one-loop effects and derive the constraints on the other components of $\lambda''$. The present upper bounds on $\lambda''$ are $\mathcal{O}(1)$ for sfermion mass $\bar{m} \sim 100$ GeV.

Fig. 1. Tree diagram for the proton decay into light gravitino or axino.
except those on $\lambda'_{112}$ and $\lambda''_{113}$ [12]. As we will see, the derived upper bounds on $\lambda'_{ijk}$ in gauge-mediated SUSY breaking models are much stronger than the currently existing bounds for a wide range of $m_{3/2}$ and $F_a$.

At low energy scales $\sim 1$ GeV where all massive particles are integrated out, the proton decay $p \to \psi +$ light meson ($\psi =$ gravitino or axino) can be described by an effective 4-fermion operator, $\mathcal{E}_{\text{eff}} = u\bar{d}s\psi$ or $u\bar{d}d\psi$, in the quark mass eigenstate basis. (See [5] for the detailed kinematic structure of these 4-fermion operators, which is not essential for our discussion in this paper.) At tree approximation, only $\lambda'_{112}$ can produce such an effective operator (see Fig. 1), thereby is constrained as $4$. In order for the other $\lambda'_{ijk}$ to produce the flavor structure $uds$ or $udd$, it must be supplemented by flavor changing interactions in the model, which is possible at one-loop order. For instance, $\lambda''_{112}$ and $\lambda''_{113}$ can induce $\mathcal{E}_{\text{eff}}$ once they are combined with the flavor change $b \to d$, while $\lambda''_{112}$ and $\lambda''_{113}$ can do it with the flavor changes $c \to u$ and $i \to u$, respectively. The other four couplings need flavor changes in order to lead to a proton decay, e.g. $(c,b) \to (u,d)$ or $(u,s)$ for $\lambda''_{113}$ and $\lambda''_{223}$, $(i,b) \to (u,d)$ or $(u,s)$ for $\lambda''_{113}$ and $\lambda''_{223}$.

To proceed, let us collect the couplings which are relevant for the proton decay into light gravitino or axino at one-loop order. First of all, one needs the standard CKM matrix element, $\tan \beta$, where all massive particles are integrated out, the proton decay induced by $\lambda''_{ijk}$. In gauge-mediated SUSY breaking models, flavor changing neutral current interactions in $R_p$-conserving sector are highly suppressed. Then the necessary flavor change takes place through the exchange of $W^\pm$, charged Higgs or charginos. The flavor changing interactions which we will use in the subsequent discussions include the $W$-boson coupling:

$$-V_{ij}\frac{g}{\sqrt{2}} W^\mu \bar{u} \gamma^\mu P_L d^j + \text{H.c.},$$

and the charged Higgs boson coupling:

$$V_{ij} H^+ \left[ f^{(d)}_j \tan \beta \bar{n}^i P_R d^j + f^{(u)}_i \cot \beta \bar{n}^i P_L d^j \right] + \text{H.c.},$$

where $g$ is the $SU(2)$ gauge coupling, $V_{ij}$ is the CKM matrix element, $\tan \beta$ is the ratio of Higgs vacuum expectation values, and $f^{(u,d)}_i$ denote the quark Yukawa couplings, i.e.

$$f^{(u)}_i = \frac{g m^{(u)}_i}{\sqrt{2} m_w \sin \beta}, \quad f^{(d)}_j = \frac{g m^{(d)}_j}{\sqrt{2} m_w \cos \beta}. \quad (9)$$
Fig. 2. One loop diagrams for the proton decay into light gravitino or axino.

(Here \(m_W\), \(m^{(\mu)}\) and \(m^{(\mu)}\) denote the masses of W-boson, up- and down-type quarks respectively.) In addition to these, we will use the following chargino-quark-squark interactions also:

\[
V_{ij}^{*} \left( \overline{\chi^{+}_{i}} \left( -g \cos \phi_{L} P_{L} + f_{ij}^{(\mu)} \sin \phi_{R} P_{R} \right) u' \tilde{d}_{L}^{+} \right.
+ f_{ij}^{(d)} \sin \phi_{R} \overline{\chi^{+}_{i}} P_{L} u' \tilde{d}_{L}^{+}
+ \frac{1}{g} \overline{\chi^{+}_{i}} \left( g \sin \phi_{L} P_{L} + f_{ij}^{(\mu)} \epsilon_{R} \cos \phi_{R} P_{R} \right) u' \tilde{d}_{L}^{+}

+ f_{ij}^{(d)} \cos \phi_{L} \overline{\chi^{+}_{i}} P_{L} u' \tilde{d}_{L}^{+}

+ V_{ij} \overline{\chi_{i}} \left( -g \cos \phi_{R} P_{L} + f_{ij}^{(d)} \sin \phi_{L} P_{R} \right) d' \tilde{u}_{L}^{+}

+ f_{ij}^{(d)} \sin \phi_{R} \overline{\chi_{i}} P_{L} d' \tilde{u}_{L}^{+}

+ \frac{1}{g} \overline{\chi_{i}} \left( g \sin \phi_{R} P_{L} + f_{ij}^{(d)} \cos \phi_{L} P_{R} \right) d' \tilde{u}_{L}^{+}

+ f_{ij}^{(d)} \epsilon_{R} \cos \phi_{R} \overline{\chi_{i}} P_{L} d' \tilde{u}_{L}^{+} \right) + \text{H.c.},
\]

where \(\epsilon_{R} = \text{sign}(\mu M_{L} - m_{h}^{(\mu)} \sin 2\beta)\) for the gaugino mass \(M_{2}\) and the Higgsino mass parameter \(\mu\). The chargino mixing angles \(\phi_{L,R}\) are given by

\[
\tan 2\phi_{L} = \frac{2\sqrt{2} m_{W} (M_{2} \cos \beta + \mu \sin \beta)}{M_{2}^{2} - m_{h}^{2} \cos 2\beta},
\]

\[
\tan 2\phi_{R} = \frac{2\sqrt{2} m_{W} (M_{2} \sin \beta + \mu \cos \beta)}{M_{2}^{2} - m_{h}^{2} \cos 2\beta}.
\]

All one loop diagrams which trigger a proton decay by having a \(\chi^{+}\)-vertex can be divided into the following three categories; (a) diagrams with radiative corrections to the \(\chi^{+}\)-vertex (Fig. 2), (b) box diagrams (Fig. 2), (c) diagrams with radiative corrections to the gravitino or axino vertex (Fig. 3). Relative to the tree diagram of Fig. 1, one loop diagrams involving \(\chi_{ij}^{0}\) will be suppressed by the factor \(\xi_{ij}^{0}\), more explicitly

\[
\frac{A_{\text{loop}}^{ij}}{A_{\text{tree}}^{ij}} = \xi_{ij}^{0} \frac{A_{\text{tree}}^{ij}}{A_{112}^{ij}},
\]

where
where $A_{\text{tree}}$ denotes the tree amplitude of Fig. 1, while $A_{\text{loop}}$ stand for the loop amplitudes of Fig. 2 and Fig. 3 which involve the insertion of $\lambda'^\nu_{ijk}$. The upper bounds on $\lambda'^\nu_{ijk}$ resulting from those one loop diagrams can be easily read off from (4) by taking into account the suppression factor $\xi_{ijk}$:

$$\mathcal{X}_{ijk} \leq 5 \times 10^{-16} \left( \frac{1}{\xi_{ijk}} \right) \left( \frac{\bar{m}}{300 \, \text{GeV}} \right)^2 \left( \frac{m_{3/2}}{1 \, \text{eV}} \right)$$

$$\mathcal{X}_{ijk} \leq 7 \times 10^{-16} \left( \frac{1}{\xi_{ijk}} \right) \left( \frac{\bar{m}}{300 \, \text{GeV}} \right)^2 \left( \frac{F_d}{10^{10} \, \text{GeV}} \right)$$

$$\times \left( \frac{1}{\epsilon_q} \right).$$

(13)

In the following, we will estimate the size of $\xi_{ijk}$ for the loop diagrams depicted in Fig. 2 and Fig. 3. Let us first consider the type (a) and (b) diagrams in Fig. 2. It turns out that type (a) diagrams (with the charged Higgs exchange) dominate in this case. The resulting suppression factors are given by

$$\xi_{ijp} \approx \frac{1}{(4\pi)^2} f_{ij}^{(a)} f_{pq}^{(b)} V_{1j} V_{1p}$$

$$= \frac{g^2}{16\pi^2} \frac{1}{m_W^2 \sin(2\beta)} m_{ij} m_{pq} V_{1j} V_{1p},$$

(14)

where $(p,q) = (1,1), (1,2)$ or $(2,1)$. It is worth noting that these suppression factors are rather insensitive to the details of unknown superparticle masses.

Although it depends more on the details of superparticle spectrum, for $\lambda'_{113}$ and $\lambda'_{123}$, one can get a much stronger bound through the diagrams in Fig. 3. For instance, we find that the loop suppression factors of Fig. 3(i) are given by

$$\xi_{123} = \frac{g^2}{16\pi^2} V_{31} V_{33} \frac{m_b}{m_W} \frac{m^2_{ij}}{\bar{m}^2} \frac{\sin 2\phi_L}{\sin 2\phi_L}$$

$$= 5 \times 10^{-7} \left( \frac{\delta m^2_{ij}}{m^2} \right) \left( \frac{300 \, \text{GeV}}{\bar{m}} \right)^2 \sin 2\phi_L.$$

$$\xi_{113} = \frac{g^2}{16\pi^2} V_{32} V_{33} \frac{m_b}{m_W} \frac{m^2_{ij}}{\bar{m}^2} \frac{\sin 2\phi_L}{\sin 2\phi_L}$$

$$= 2 \times 10^{-6} \left( \frac{\delta m^2_{ij}}{m^2} \right) \left( \frac{300 \, \text{GeV}}{\bar{m}} \right)^2 \sin 2\phi_L,$$

(15)

where we have assumed that all superparticle masses including the chargino masses are approximately same as the universal squark mass $\bar{m}$, and $\delta m^2_{ij}$ is...
the difference between the two chargino mass-squared,
\[ \delta m^2_\chi = |m^2_\chi - m^2_\tilde{g}|. \] (16)

Here the extra \( m_f \)-dependence is due to the GIM-cancellation. In fact, one can consider a diagram which is similar to Fig. 3(i) but including the insertion of the left-right squark instead of the chargino, however here we will ignore this extra complication.

For the axino case, the correct suppression factor of Fig. 3(ii) is
\[ \delta m^2_\tilde{a} = |m^2_\tilde{a} - m^2_\tilde{g}|. \]

As a result, for the axino case since it involves the insertion of the left-right squark (which is similar to Fig. 3(i) but including the insertion of the left-right squark) the amplitude of such diagram is heavily suppressed by the GIM mechanism, and thus it does not give a bound on \( \lambda''_{112} \) or \( \lambda''_{113} \). However the amplitude for Fig. 3(ii) in the axino case since it involves the insertion of the left-right squark
\[ \delta m^2_\chi = |m^2_\chi - m^2_\tilde{g}|. \]

This extra factor will lead to one order of magnitude stronger bound for hadronic-type axino, however here we will ignore this extra complication.

If the charginos are degenerate or the chargino mixing \( |\sin 2\phi_1| \ll 1 \), the bounds from Fig. 3(i) will be significantly weakened. In this case, the dominant contribution would come from Fig. 3(ii) or 3(iii) which involves the insertion of the left-right squark mixing:
\[ m^2_\chi \mu \tan \tilde{\beta} \tilde{d}_R \tilde{d}_R + m^2_\chi \mu \cot \tilde{\beta} \tilde{u}_L \tilde{u}_R + \text{H.c.} \] (17)

The corresponding loop suppression factors are given by
\[ \xi_{123} = \frac{g^2}{8\pi^2} V_{31} V_{33} \frac{\mu m_h}{m^2} \frac{m^2_\chi}{m^2_\tilde{g}} \approx 4 \times 10^{-7} \left( \frac{\mu}{m} \right) \left( \frac{300 \text{ GeV}}{m} \right)^3, \]
\[ \xi_{113} = \frac{g^2}{8\pi^2} V_{32} V_{33} \frac{\mu m_h}{m^2} \frac{m^2_\chi}{m^2_\tilde{g}} \approx 1 \times 10^{-6} \left( \frac{\mu}{m} \right) \left( \frac{300 \text{ GeV}}{m} \right)^3, \] (18)

where again it is assumed that all superparticle masses are the approximately same as the universal squark mass \( m_\tilde{g} \). In fact, there arises an extra complication for Fig. 3(ii) in the axino case since it involves the axino coupling to gauge multiplets \( e_+ / 8\pi^2 \) in Eq. (5), while the tree diagram of Fig. 1 involves only the axino coupling to the light quark multiplets \( e_+ / 8\pi^2 \) in Eq. (5) for \( I = q \) where \( q = (u, d, s) \) stands for the light quark multiplets). As a result, for the axino case, the correct suppression factor of Fig. 3(ii) is obtained by multiplying the factor \( e_+ / 8\pi^2 \) to the results of Eq. (18). This extra factor will lead to one order of magnitude stronger bound for hadronic-type axino, however here we will ignore this extra complication.

### Table 1

<table>
<thead>
<tr>
<th>Coupling</th>
<th>Upper Bound I</th>
<th>Upper Bound II</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda''_{112} )</td>
<td>( 5 \times 10^{-9} )</td>
<td>( 7 \times 10^{-10} )</td>
</tr>
<tr>
<td>( \lambda''_{113} )</td>
<td>( 3 \times 10^{-9} )</td>
<td>( 7 \times 10^{-10} )</td>
</tr>
<tr>
<td>( \lambda''_{123} )</td>
<td>( 1 \times 10^{-7} )</td>
<td>( 2 \times 10^{-8} )</td>
</tr>
<tr>
<td>( \lambda''_{112} )</td>
<td>( 5 \times 10^{-8} )</td>
<td>( 7 \times 10^{-9} )</td>
</tr>
<tr>
<td>( \lambda''_{113} )</td>
<td>( 1 \times 10^{-8} )</td>
<td>( 7 \times 10^{-9} )</td>
</tr>
<tr>
<td>( \lambda''_{123} )</td>
<td>( 5 \times 10^{-9} )</td>
<td>( 7 \times 10^{-9} )</td>
</tr>
</tbody>
</table>

Applying the loop suppression factors of Eqs. (14), (15) and (18) to Eq. (13), one can easily derive the upper bounds on \( \lambda''_{ijk} \). We summarize the numerical results in Table 1. In deriving these, we take \( \sin 2\beta = 1 \) in (14) and ignored the contribution from Fig. 3(ii) for the axino case, which would lead to conservative results. We also assumed that all superparticle masses are the approximately same as the universal squark mass \( m_\tilde{g} \), and also \( \mu \approx m_\tilde{g} \). For the numerical values of the quark masses, CKM matrix elements and etc., the values in Ref. [15] are used.

To conclude, we have examined the proton decay involving a light gravitino or axino in gauge-mediated supersymmetry breaking models to derive constraints on the \( R \) parity and baryon number violating Yukawa couplings \( \lambda''_{ijk} \). Considering the decay amplitudes at one-loop order, we could get upper bounds on all of those couplings. The results summarized in Table 1 show that, for a wide range of the gravitino mass \( m_{3/2} \), the bounds on all \( \lambda''_{ijk} \) are much stronger than the currently existing bounds. The bounds on \( \lambda''_{113} \) and \( \lambda''_{123} \) are particularly strong due to the contributions from Fig. 3. In supersymmetric models with \( \ell(1/2) \), if axino is lighter than the proton, all \( \lambda''_{ijk} \) are similarly constrained by the proton decay into light axino.

We thank E.J. Chun for helpful discussions. This work was supported in part by KOSEF through CTP.
of SNU, Basic Science Research Institutes Program BSRI-97-2434 and KAIST Basic Science Research Program.

References


Quasi degenerate neutrino masses with universal strength Yukawa couplings

G.C. Branco 1, M.N. Rebelo 2, J.I. Silva-Marcos 3

CFIF / IST and Departamento de Física, Instituto Superior Técnico, Av. Rovisco Pais, P-1096 Lisboa Codex, Portugal

Received 10 February 1998
Editor: R. Gatto

Abstract

A simple ansatz is proposed for neutrino and charged lepton mass matrices, within the framework of universal strength for Yukawa couplings. In this framework all Yukawa couplings have equal moduli and the flavour dependence is only in their phases. We take into account the solar neutrino deficit and the atmospheric neutrino anomaly, assuming three neutrino families only. The ansatz leads in a natural way to small mixing involving neutrinos of quasi degenerate masses, as required to explain the solar neutrino deficit in the non-adiabatic MSW solution, while having the large mixing necessary to explain the atmospheric neutrino anomaly. © 1998 Published by Elsevier Science B.V. All rights reserved.

1. Introduction

Presently the Standard Model (SM) enjoys quite a remarkable success when confronted with experiment. However, there is recent experimental evidence pointing towards physics beyond the SM in the leptonic sector, to wit, the solar neutrino deficit, the atmospheric neutrino problem, and the results of the LSND collaboration suggesting that neutrino oscillations might have been observed in an accelerator experiment.

The solar neutrino data obtained by several different experiments [1], indicate a deficit in the number of observed neutrinos by comparison to the Standard Solar Model (SSM) predictions for the solar neutrino fluxes [2]. The solar neutrino deficit is explained in terms of oscillations of the electron neutrino into some other neutrino species. In the framework of the MSW mechanism [3] there are two sets of solutions, the adiabatic branch (AMSW) requiring a large mixing (sin$^2\theta = 0.65\pm0.85$) [4] and the non adiabatic branch (NAMSW) requiring a small mixing (sin$^2\theta = (0.1\pm2) \times 10^{-2}$) [4] with $\Delta m^2 = (0.3\pm1.2) \times 10^{-5}$ eV$^2$. The small mixing solution seems to be favoured by the present data. In the framework of vacuum oscillation only three separate regions within a narrow range of parameters ($\Delta m^2 = 5\pm8 \times 10^{-11}$ eV$^2$ and $\sin^22\theta = 0.65\pm1$) are allowed [4].

Several experiments have measured the ratio of the number of muon neutrinos by the number of electron neutrinos produced in the atmosphere through the decay of pions and kaons with subse-
quent decay of secondary muons [5]. The combined results lead to:

\[
R = \frac{\left( n_{\mu}/n_{\nu_e}\right)_{\text{Data}}}{\left( n_{\mu}/n_{\nu_e}\right)_{\text{SM}}} \approx 0.6
\]  

where \( \left( n_{\mu}/n_{\nu_e}\right)_{\text{Data}} \) is the measured ratio of muon-neutrino to electron-neutrino events, while \( \left( n_{\mu}/n_{\nu_e}\right)_{\text{SM}} \) is the expected ratio assuming no oscillations. This anomaly can be caused by oscillations of the atmospheric muon neutrinos into another type of neutrino with large mixing angle \( \sin^2 2\theta = 0.6-1.0, \Delta m^2 = (0.3-3) \times 10^{-2} \text{eV}^2 \).

The LSND group has reported evidence for \( \nu_\mu - \nu_\tau \) and \( \nu_\mu - \nu_e \) oscillations [6]. However no other accelerator experiment confirmed their result thus strongly reducing the allowed parameter space. The resulting experimental constraints on neutrino masses and mixings are such that no combined solution to the solar, atmospheric and LSND results can be found in the framework of three neutrinos only, since the required neutrino mass differences do not satisfy the relation \( \Delta m^2_{12}, \Delta m^2_{23} = \Delta m^2_{13} \). We have chosen to consider simply three neutrino families without additional sterile neutrinos and not to take into consideration the LSND data.

Astrophysical considerations, in particular the possibility that neutrinos constitute the hot dark matter, favour neutrino masses of the order of a few eV [7], which combined with the above constraints leads to a set of highly degenerate neutrinos.

In this paper we propose a simple ansatz within the framework of universal strength for Yukawa couplings (USY) [8] which leads to quasi-degeneracy of neutrino masses and provides a solution to the solar and atmospheric neutrino problems. In USY all Yukawa couplings have equal moduli so that the flavour dependence is only contained in their phases. For the quark sector it has been shown [9,10] that the USY hypothesis leads to a highly predictive and successful ansatz for quark masses and mixings. Various mixing schemes for the leptonic sector have been suggested [11]. The possibility of quasi-degenerate neutrino masses has been recently considered in the literature in the context of some specific symmetry or ansatz. An important feature of our ansatz lies on the fact that it can accommodate in a natural way small mixing involving a set of highly degenerate neutrinos in the MSW solution to the solar neutrino problem while also having the large mixing necessary to explain the atmospheric neutrino data.

2. Degeneracy in USY

The USY hypothesis leads to neutrino mass matrices of the form:

\[
M_\nu = c_\nu \left[ e^{i\theta_{ij}} \right]
\]  

where \( c_\nu \) is an overall constant. Let us derive the conditions which should be satisfied so that the matrices of Eq. (2) lead to at least two degenerate neutrinos. It is useful to introduce the dimensionless Hermitian matrix \( H_\nu = M_\nu M_\nu^\dagger / 3c_\nu^2 \) which can be written as:

\[
H_\nu = \begin{pmatrix}
1 & r_1e^{i\varphi_1} & r_2e^{i\varphi_2} \\
r_1e^{-i\varphi_1} & 1 & r_3e^{i\varphi_3} \\
r_2e^{-i\varphi_2} & r_3e^{-i\varphi_3} & 1
\end{pmatrix}
\]  

where the off diagonal elements \( r_i e^{i\varphi_i} \) are the sum of products of pure phase elements of \( M_\nu \):

\[
(H_\nu)_{ij} = r_i e^{i\varphi_i} = \frac{1}{3} \left[ e^{i(\theta_{ij} - \theta_{11})} + e^{i(\theta_{ij} - \theta_{22})} + e^{i(\theta_{ij} - \theta_{33})} \right]
\]  

with analogous expressions for \( (H_\nu)_{13} \) and \( (H_\nu)_{23} \). It can be readily verified that if at least two of the eigenvalues of \( H_\nu \) are equal, then the following relation holds:

\[
\left[ 1 - \chi^2 \right] = \left[ 1 - \frac{\chi - \delta}{2} \right]^2
\]  

where:

\[
\delta = \det (H_\nu) = \lambda_1 \lambda_2 \lambda_3,
\]

\[
\chi \equiv \chi (H_\nu) = \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3
\]  

and the \( \lambda_i = 3m_i^2 / (m_1^2 + m_2^2 + m_3^2) \) denote the eigen-
values of $H_v$. In the derivation of Eq. (5) we took into account that $tr(H_v) = \lambda_1 + \lambda_2 + \lambda_3 = 3$. The invariants $\chi$ and $\delta$ can be expressed in terms of $r_i$ and $\phi$, with all generality, as:

$$\chi = 3 - r_1^2 - r_2^2 - r_3^2,$$

$$\delta = 1 + 2r_1r_2r_3\cos(\varphi_1 + \varphi_3 - \varphi_2) - r_1^2 - r_2^2 - r_3^2$$

(7)

The following ‘‘spherical’’ parametrization for the $r_i$ is useful:

$$r_1 = \left[3(1 - \chi/3)\right]^{1/2} \sin \theta \cos \phi,$$

$$r_2 = \left[3(1 - \chi/3)\right]^{1/2} \sin \theta \sin \phi,$$

$$r_3 = \left[3(1 - \chi/3)\right]^{1/2} \cos \theta$$

(8)

with $0 \leq \theta, \phi \leq \pi$. From Eqs. (7) and (8), one obtains:

$$\sin^2 \theta \cos \theta \cdot \sin(2\phi) \cdot \cos(\varphi_1 + \varphi_3 - \varphi_2)$$

$$= \frac{2}{3\sqrt{3}} \left[ 1 - \frac{\chi - \delta}{2} \right] \left[ 1 - \frac{\chi}{3} \right]^{3/2}$$

(9)

In the case of at least two degenerate neutrinos Eq. (9) can be combined with Eq. (5) leading to the condition:

$$\sin^2 \theta \cos \theta \cdot \sin(2\phi) \cdot \cos(\varphi_1 + \varphi_3 - \varphi_2) = \frac{2}{3\sqrt{3}}$$

(10)

which can only be satisfied for $\cos \theta = 1/\sqrt{3}$, $\sin(2\phi) = 1$ and $\cos(\varphi_1 + \varphi_3 - \varphi_2) = 1$. Therefore, the necessary and sufficient conditions for $H_v$ to have two degenerate eigenvalues are:

$$\varphi_1 + \varphi_3 - \varphi_2 = 0 \mod 2\pi, \quad r_1 = r_2 = r_3$$

(11)

Within the USY framework, it can be shown that for matrices $M_v$ having at least two degenerate eigenvalues, there is a weak-basis where $M_v$ has one of the following forms (modulo trivial permutations):

$$M_v^I = c_v K \cdot \begin{bmatrix} e^{i\alpha} & 1 & 1 \\ 1 & e^{i\alpha} & 1 \\ 1 & 1 & e^{i\alpha} \end{bmatrix};$$

$$M_v^{II} = c_v K \cdot \begin{bmatrix} e^{i\alpha} & 1 & 1 \\ 1 & e^{-2i\alpha} & 1 \\ 1 & 1 & 1 \end{bmatrix};$$

$$M_v^{III} = c_v K \cdot \begin{bmatrix} e^{i\alpha} & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

(12)

where $K = \text{diag}(e^{i\phi_1}, e^{i\phi_2}, e^{i\phi_3})$ and $c_v$ is a real constant.

The first two cases of Eq. (12) are of special interest since for $\alpha = 2\pi/3$ they lead to three degenerate neutrino masses, while for $\alpha \neq 2\pi/3$ they lead to only two degenerate mass eigenvalues.

3. A special ansatz

In order to have an ansatz with predictions for the leptonic mixing matrix, one has to specify the structure of the charged lepton mass matrix $M_v$ together with the structure of $M_{\nu}$.

Our guiding principle is the assumption that all leptonic Yukawa couplings obey the USY hypothesis. Furthermore, we choose, within USY, the same structure of phases for the charged lepton and neutrino mass matrices. Guided by the above ideas we propose the following specific ansatz:

$$M_v = c_v \begin{bmatrix} e^{-i\alpha} & 1 & 1 \\ 1 & e^{i\alpha} & 1 \\ 1 & 1 & e^{i\beta} \end{bmatrix};$$

(13)

$$M_\nu = c_\nu \begin{bmatrix} e^{i\alpha} & 1 & 1 \\ 1 & e^{i\alpha} & 1 \\ 1 & 1 & e^{i\beta} \end{bmatrix}$$

with $c_\nu, c_v$ real constants.

The leptonic mixing matrix $V$ appearing in the charged weak current is then given by:

$$V = U_v^\dagger \cdot U_\nu$$

(14)

where the matrix $U_\nu$ diagonalizes $M_\nu M^\dagger_\nu$ and the
matrix \( U_e \), diagonalizes \( M_e M_
u^T \) in the case of Dirac neutrinos. Note that both \( M_e \) and \( M_
u \) have only three parameters each. As a result, both \( U_e \) and \( U_
u \) will be entirely fixed by charged lepton and neutrino mass ratios. Due to the observed strong hierarchy in the charged lepton mass spectrum, the parameters \((a,b)\) will be close to zero. From the form of \( M_e \) in Eq. (13) one can derive exact expressions for the phases \((a,b)\) in terms of charged lepton mass ratios. In leading order, one obtains:

\[
\frac{m_
u}{m_e} \approx \begin{pmatrix} m_{\mu} & m_{\mu} & m_{\mu} \\ m_{\mu} & m_{\mu} & m_{\mu} \\ m_{\mu} & m_{\mu} & m_{\mu} \end{pmatrix} \quad ; \quad b = \frac{9}{2} \frac{m_{\mu}}{m_e} \tag{15}
\]

In the neutrino sector, it was already pointed out that the square mass differences necessary to explain the solar neutrino data, together with the requirement that neutrinos constitute the hot dark matter, lead to highly degenerate neutrinos. The matrix \( M_\nu \) in Eq. (13) leads to threefold degeneracy for \( \alpha = \beta = 2\pi/3 \). Thus, it is useful to introduce the two small parameters \( \epsilon_2, \epsilon_3 \), defined by:

\[
\alpha = \frac{2\pi}{3} - \epsilon_2 - \epsilon_3; \quad \beta = \frac{2\pi}{3} + 2\epsilon_3 \tag{16}
\]

It is clear from this parametrization, that in the limit \( \epsilon_2, \epsilon_3 \to 0 \), one has for the phase \( \beta = -2\alpha \) (mod \( 2\pi \)), and therefore one obtains in this limit the mass matrix \( M_\nu^N \) of Eq. (12), with two exactly degenerate masses. Using Eq. (13), one calculates the eigenvalues \( \lambda_i \) of \( H_e = M_e M_e^\dagger / 3c^2_\nu \) as functions of \( \alpha \) and \( \beta \):

\[
\lambda_1 = 1 - y; \quad \lambda_2 = \frac{2 + y - \sqrt{8y^2 + y^2}}{2}; \quad \lambda_3 = \frac{2 + y + \sqrt{8y^2 + y^2}}{2} \tag{17}
\]

where \( y = (2\cos(\alpha) + 1)/3 \) and \( x = (1/3)[3 + 2\cos(\alpha) + 2\cos(\beta) + 2\cos(\alpha + \beta)]^{1/2} \). Then from Eq. (17) one can express \( \epsilon_2, \epsilon_3 \) in terms of mass ratios. In leading order, one obtains:

\[
\epsilon_{32} = \frac{1}{\sqrt{3}} \frac{\Delta m^2_{32}}{m_3^2}; \quad \epsilon_{21} = \frac{3\sqrt{3}}{4} \frac{\Delta m^2_{21}}{m_3^2} \tag{18}
\]

In order to evaluate the leptonic mixing matrix \( V \), it is convenient to start by making a weak-basis (WB) transformation defined by:

\[
H_e \equiv \frac{1}{3c^2_\nu} M_e M_e^\dagger \rightarrow H_e' = F \cdot H_e \cdot F^\dagger, \quad H_e' \equiv \frac{1}{3c^2_\nu} M_e M_e^\dagger \rightarrow H_e' = F \cdot H_e' \cdot F \tag{19}
\]

where:

\[
F = \begin{pmatrix} 1 & -1 & 1 \\ \sqrt{2} & \sqrt{6} & \sqrt{3} \\ -1 & -1 & 1 \\ \sqrt{2} & \sqrt{6} & \sqrt{3} \\ 0 & 2 & 1 \\ \sqrt{2} & \sqrt{6} & \sqrt{3} \end{pmatrix} \tag{20}
\]

Obviously, after this WB transformation the charged leptonic weak current remains diagonal, real, i.e., \( V = 1 \). This WB transformation corresponds to writing \( H_e' \) in a “heavy basis”. Due to the strong hierarchy of the charged lepton masses, in this “heavy basis” all other elements of \( H_e' \) are small, compared to the element \((3,3)\). The Hermitian matrix \( H_e' \) can then be diagonalized, and one obtains in leading order:

\[
|U_{12}^e| = \sqrt{\frac{m_\mu}{m_e}}; \quad |U_{13}^e| = \sqrt{\frac{2m_\mu}{m_e}}; \quad |U_{23}^e| = \sqrt{\frac{m_\mu}{m_e}} \tag{21}
\]

In the neutrino sector, one has a drastically different situation. After the WB transformation of Eq. (19), the first neutrino exactly decouples from the other two, i.e., \((H_e')_{12} = (H_e')_{13} = 0\).
Furthermore, the mixing between the second and third neutrino state is large and, to leading order, independent of neutrino mass ratios, one obtains:

$$U_e = \begin{bmatrix} 1 & 0 & 0 \\ \frac{\omega^* - \omega}{3} & \frac{\sqrt{2}(\omega^* - 1)}{3} & 1 - \omega \\ \frac{\sqrt{2}(\omega^* - 1)}{3} & \frac{1 - \omega}{3} & \frac{\sqrt{2}}{3} \end{bmatrix}$$  (22)

where $\omega = e^{i2\pi/3}$. In Eq. (22) the zeros are exact yet we have omitted in the other entries terms of the order $\Delta m^2_{31}/\Delta m^2_{32}$. Therefore, the leptonic mixing matrix is given, in leading order, by:

$$|V| = \begin{bmatrix} 1 & \sqrt{\frac{m_e}{m_\mu}} & \sqrt{\frac{2m_e m_\mu}{m_\tau}} \\ \sqrt{\frac{m_e}{3m_\mu}} & \frac{1}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} \\ \frac{2m_e}{3m_\mu} & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$  (23)

4. Confronting the data

Neutrino oscillations [13] have to be taken into account since neutrino experiments are discussed in terms of neutrino weak eigenstates rather than physical states. With the notation $\nu_{a(\beta)}$ for weak eigenstates and $\nu_{(a,\beta)}$ for mass eigenstates, the probability of finding $\nu_\beta$ at time $t$ having started with $\nu_a$ at $t=0$ neglecting the effect of CP violation in the leptonic sector (real $V$) is given by:

$$P(\nu_a \rightarrow \nu_\beta) = \langle \nu_\beta | \nu_\beta(t) \rangle \cdot \langle \nu_\beta | \nu_a(0) \rangle^*$$  (24)

where $E$ is the neutrino energy, $L$ is the distance travelled by the neutrino between the source and the detector, $V$ is the leptonic mixing matrix given by Eq. (14), and $\Delta m^2_{\mu}$ is defined by:

$$\Delta m^2_{\mu} = |m_\mu^2 - m_\tau^2|$$  (25)

The interpretation of the experimental data is presented in terms of two flavour mixing, in this case Eq. (24) reduces to:

$$P(\nu_a \rightarrow \nu_\beta) = \delta_{a\beta} - \sin^2 2\theta \cdot \sin^2 \left[ \frac{\Delta m^2_{\mu} L}{4E} \right]$$  (26)

hence the meaning of the experimental bounds presented in Section 1.

The translation of the bounds into the three flavour approach is quite simple for the experimental limits imposed on $\Delta m^2_{\mu}$, with $\Delta m^2_{31}$ in the solar range and $\Delta m^2_{32}$ in the atmospheric range. In the case of the atmospheric anomaly it is clear that we can disregard the term in $\sin^2(\Delta m^2_{32}/4L/E)$ and we can approximately write:

$$\sin^2 \left[ \frac{\Delta m^2_{31} L}{4E} \right] = \sin^2 \left[ \frac{\Delta m^2_{32} L}{4E} \right] \equiv S$$  (27)

so that:

$$1 - P(\nu_\mu \rightarrow \nu_e) = P(\nu_\mu \rightarrow \nu_e) + P(\nu_\mu \rightarrow \nu_\tau) = 4(V_{32}V_{32}V_{12} + V_{22}V_{22}V_{32}V_{32}) S$$  (28)

and we identify:

$$\sin^2 2\theta_{\atm} = 4(V_{32}V_{32}V_{12} + V_{22}V_{22}V_{32}V_{32})$$  (29)

In the case of the solar neutrino anomaly the range $L/E$ is such that $S$ in Eq. (27) can be averaged to $\frac{1}{2}$ and we obtain:

$$1 - P(\nu_e \rightarrow \nu_e) = 4(V_{11}V_{11}V_{31}V_{31} + V_{12}V_{12}V_{31}V_{31}) \cdot \frac{1}{2}$$

$$+ 4V_{11}V_{11}V_{21}V_{21} \sin^2 \left[ \frac{\Delta m^2_{32} L}{4E} \right]$$  (30)

in our numerical example the first term of this equation is small so, as a result, we can identify:

$$\sin^2 2\theta_{\sun} = 4V_{11}V_{11}V_{21}V_{21}$$  (31)

Our ansatz can perfectly fit the experimental bounds as is shown by the following example.
We choose as input the masses for the charged leptons:
\[ m_e = 0.511 \text{ MeV}, \quad m_{\mu} = 105.7 \text{ MeV}, \]
\[ m_{\tau} = 1777 \text{ MeV} \] (32)
which correspond to the phases \( a = 0.0214 \) and \( b = 0.2662 \) of Eq. (13). For the neutrino sector, we choose:
\[ m_{\nu_1} = 2 \text{ eV}, \]
\[ \Delta m^2_{1} = 9.2 \times 10^{-6} \text{ eV}^2, \]
\[ \Delta m^2_{2} = 5.0 \times 10^{-3} \text{ eV}^2 \] (33)
This fixes the values of the parameters \( e_1 = 3 \times 10^{-6} \) and \( e_2 = 7.2 \times 10^{-4} \). With the above input we obtain for the leptonic mixing matrix \( V \), without approximations:
\[
[V] = \begin{bmatrix}
0.9976 & 0.0692 & 0.0058 \\
0.0463 & 0.6068 & 0.7935 \\
0.0518 & 0.7918 & 0.6085 
\end{bmatrix}
\] (34)
Making use of Eqs. (29) and (31), Eq. (34) translates into:
\[ \sin^2 2\theta_{\text{atm}} = 0.935 \] (35)
and:
\[ \sin^2 2\theta_{\text{sol}} = 0.0085 \] (36)
These results are in agreement with the present experimental data. In this example the three physical neutrinos under consideration all have masses close to 2 eV. Note that the leptonic mixing matrix \( V \), which differs significantly from the observed quark mixing matrix, was obtained having as input the charged lepton and neutrino masses, with no further parameters.

5. Concluding remarks

We have suggested an ansatz for the neutrino and charged lepton mass matrices, within the framework of universality of strength of Yukawa couplings. The ansatz has the same structure of phases for the neutrino and charged lepton mass matrices, with the only non-vanishing phases along the diagonal. Both \( M_\nu \) and \( M_e \) have only three parameters each, two phases and an overall real constant. These parameters are completely fixed by the value of charged lepton and neutrino masses, which implies that the ansatz is highly predictive, with full calculability of the leptonic mixing matrix, i.e., \( V \) is completely fixed by charged lepton and neutrino mass ratios. We have shown that the ansatz naturally leads on the one hand to small mixing among \( \nu_e \) and \( \nu_\mu \), \( \nu_\tau \), associated to \( \Delta m^2_{1} \sim 10^{-3} \text{ eV}^2 \) thus explaining the solar neutrino deficit in the non-adiabatic MSW solution, and on the other hand leads to large mixing between \( \nu_\mu \) and \( \nu_\tau \), as required to explain the atmospheric neutrino anomaly.

We find our results specially appealing, since one has an unified view of all Yukawa couplings, i.e., both quark and lepton Yukawa couplings have universal strength, with the flavour dependence being all contained in their phases. Furthermore, both in the lepton and quark sectors [10] one has simple ansätze within USY, whose distinctive feature is having a number of independent parameters equal to the number of elementary fermion masses. As a result, one has highly predictive schemes, with the fermion mixings expressed in terms of fermion mass ratios with no free parameters.

Acknowledgements

We would like to thank M. Tanimoto for enlightening discussions during the preparation of this work.

References


Quantization of the nonlinear Schrödinger equation on the half line

Mario Gattobigio a, Antonio Liguori b, Mihail Mintchev a

a Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Dipartimento di Fisica dell’Università di Pisa, Piazza Torricelli 2, 56100 Pisa, Italy
b International School for Advanced Studies, 34014 Trieste, Italy

Received 2 February 1998
Editor: L. Alvarez-Gaumé

Abstract

We establish the second quantized solution of the nonlinear Schrödinger equation on the half line with a mixed boundary condition. The solution is based on a new algebraic structure, which we call boundary exchange algebra and which substitutes, in the presence of boundaries, the familiar Zamolodchikov-Faddeev algebra. © 1998 Elsevier Science B.V. All rights reserved.

PACS: 03.70.+k; 11.10.Kk
Keywords: Integrable systems; Quantization; Boundary effects

The recent interest in quantum field theory on the half line \( \mathbb{R}_+ = \{ x \in \mathbb{R} : x > 0 \} \) is related to some successful applications in open string theory, dissipative quantum mechanics and impurity problems in condensed matter physics. This paper concerns the quantization of the nonlinear Schrödinger equation (NLS)

\[
(i \partial_x + \partial_x^2) \Phi(t, x) = 2g |\Phi(t, x)|^2 \Phi(t, x),
\]

\[ g > 0, \]

on \( \mathbb{R}_+ \) with the boundary condition

\[
\lim_{x \downarrow 0} (\partial_x - \eta) \Phi(t, x) = 0, \quad \eta \geq 0.
\]

Our main result is the exact second quantized solution of the boundary valued problem (1), (2).

When considered on the whole line \( (x \in \mathbb{R}) \), Eq. (1) gives rise to one of the most popular and extensively studied integrable systems, which has been solved [1–4] by the method of inverse scattering transform. Let us summarize briefly those main results in \( \mathbb{R} \), which turn out to be relevant for our investigation in \( \mathbb{R}_+ \). A convenient starting point is a classical solution of Eq. (1), discovered time ago by Rosales [5] and reading

\[
\Phi(t, x) = \sum_{n=0}^{\infty} (-g)^n \Phi^{(n)}(t, x),
\]
\[ \Phi^{(n)}(t,x) = \int \prod_{i=1}^{n} \frac{dp_i}{2\pi} \frac{dq_i}{2\pi} e^{i(x_i - t_i)q_i - \frac{i\epsilon}{2}\sum_{i=1}^{n} \sum_{j=1}^{n} (x_j - t_j)(x_i - t_i)} \prod_{i=1}^{n} \left[ (p_i - q_{i-1}) (p_i - q_i) \right] . \]  

The bar indicates complex conjugation and the integration is defined by the principal value prescription. It is assumed that the function \( \lambda \) is such that the integrals in (4) exist and the series (3) converges for sufficiently small \( \epsilon \). Any Schwartz test function meets for these requirements. A remarkable property of the solution (3), (4) is that its general structure is preserved by the quantization. Indeed, following [2–4], the quantum solution on \( \mathbb{R} \) admits the series expansion (3) and the \( n \)-th order contribution has a similar form, namely

\[ \Phi^{(n)}(t,x) 
= \int \prod_{i=1}^{n} \frac{dp_i}{2\pi} \frac{dq_i}{2\pi} e^{i(x_i - t_i)q_i - \frac{i\epsilon}{2}\sum_{i=1}^{n} \sum_{j=1}^{n} (x_j - t_j)(x_i - t_i)} \prod_{i=1}^{n} \left[ (p_i - q_{i-1}) (p_i - q_i) \right] . \]

Besides the \( \epsilon \) prescription to contour poles, the new and fundamental feature is the substitution of the the functions \( \{ \lambda(k), \bar{\lambda}(k) \} \) with the generators \( \{ a(k), a^\dagger(k) \} \) of the Zamolodchikov-Faddeev (ZF) [6] algebra \( \mathcal{A}_R \) defined by

\[ a(k_1) a(k_2) = R(k_2,k_1) a(k_2) a(k_1) = 0 , 
\]

\[ a^\dagger(k_1) a^\dagger(k_2) = R(k_2,k_1) a^\dagger(k_2) a^\dagger(k_1) = 0 , 
\]

\[ a(k_1) a^\dagger(k_2) - R(k_2,k_1) a^\dagger(k_2) a(k_1) = 0 , 
\]

\[ a^\dagger(k_1) a^\dagger(k_2) - R(k_2,k_1) a^\dagger(k_2) a(k_1) = 2\pi \delta(k_1 - k_2) , \]

where

\[ R(p,k) = \frac{p-k-ig}{p-k+ig} . \]

In Eq. (5) \( \{ a(k), a^\dagger(k) \} \) are taken in the Fock representation (see e.g. [7,8]) \( \mathcal{F}_R \) of \( \mathcal{A}_R \), which provides therefore the state space of the quantum NLS. As usual, the \( * \)-operation is realized as Hermitian conjugation with respect to the scalar product in \( \mathcal{F}_R \). In order to give a precise meaning of the cubic term in (1), we introduce normal ordered \( : \ldots: \) products involving \( \Phi \) and \( \Phi^* \). In such a product, all creation operators \( a^\dagger \) stand to the left of all annihilation operators \( a \). In view of Eq. (6), one must further specify the ordering of creators and annihilators themselves. We define \( : \ldots: \) to preserve the original order of the creators; the original order of two annihilators is preserved if both of them belong to the same \( \Phi \) or \( \Phi^* \) and inverted otherwise. The quantum version of Eq. (1) is then obtained by the substitution

\[ \Phi(t,x) \rightarrow : \Phi \Phi^* \Phi : (t,x) , \]

and can be verified explicitly [4]. An essential point is also that the equal-time canonical commutation relations

\[ [\Phi(t,x), \Phi(t,y)] = [\Phi^*(t,x), \Phi^*(t,y)] \]

\[ = 0 , \]

\[ [\Phi(t,x), \Phi^*(t,y)] = \delta(x-y) , \]

hold [4] on a dense domain in \( \mathcal{F}_R \).

Eq. (5) is the quantum inverse scattering transform for the NLS on \( \mathbb{R} \). It allows to reconstruct the off-shell quantum field \( \Phi(t,x) \) from the scattering data encoded in the Fock representation \( \mathcal{F}_R \) of \( \mathcal{A}_R \). One can show [1–4] in fact, that the relative asymptotic states can be represented by

\[ |k_1, \ldots, k_n \rangle_{\text{in}} = a^\dagger(k_1) \ldots a^\dagger(k_n) \Omega , \]

\[ k_1 > \ldots > k_n , \]

\[ |p_1, \ldots, p_n \rangle_{\text{out}} = a^\dagger(p_1) \ldots a^\dagger(p_n) \Omega , \]

\[ p_1 < \ldots < p_n , \]

where \( \Omega \in \mathcal{F}_R \) is the vacuum state and \( \{ k_1, \ldots, k_n \} \) and \( \{ p_1, \ldots, p_n \} \) denote the momenta of the incoming and outgoing particles respectively. Varying \( n \geq 0 \), the linear subspaces generated by the vectors (11) and (12) give rise to the corresponding asymptotic spaces \( \mathcal{F}_R^{\text{in}} \) and \( \mathcal{F}_R^{\text{out}} \), which turn out [8] to be separately dense in \( \mathcal{F}_R \). By means of Eqs. (6) one
easily derives the scattering amplitudes, which have the factorized form
\[
\text{out} \langle p_1,\ldots,p_n|k_1,\ldots,k_n|\text{in} = \delta_{mn} \prod_{i=1}^{n} 2\pi\delta(p_i - k_{n+1-i}) \times \prod_{i,j=1, i < j} R(p_i, p_j) .
\]  
As it can be expected from integrability, these amplitudes vanish unless \( n = m \). Moreover, the particle momenta are separately conserved, \( p_i = k_{n+1-i}, \ i = 1,\ldots,n \). From Eq. (13) we deduce also that the exchange factor \( R \) in the ZF algebra \( \mathcal{B}_R \) is actually the two-body scattering matrix.

At this stage we have enough background for facing the boundary valued problem (1), (2) on the half line \( \mathbb{R}_+ \). Its classical integrability has been investigated in [9]. After some algebra one can verify [9,10] that besides Eq. (1), the field \( \Phi(t,x) \) defined by Eqs. (3), (4) solves also the boundary condition (2), provided that the function \( \lambda(k) \) satisfies the reflection condition
\[
\lambda(k) = B(k) \lambda(-k) ,
\]  
with reflection coefficient given by
\[
B(k) = \frac{k - i\eta}{k + i\eta} .
\]  
So, as a first tentative to quantize Eqs. (1), (2), one may try to keep Eqs. (3), (5) and to implement in \( \mathcal{B}_R \) the analog of condition (14), namely \( a(k) = B(k)a(-k) \). It is easily seen however that such a constraint is not compatible with the exchange relations (6) and the two-body bulk scattering matrix (7).

The next natural conjecture at this point is that the quantum solution of Eqs. (1), (2) is still of the form (3), (5), but for the replacement of \( \mathcal{A}_R \) with another appropriate algebraic structure \( \mathcal{B}_R \), called in what follows boundary algebra. It will be shown below that this conjecture is right and the main problem is to determine \( \mathcal{B}_R \), finding in particular the quantum counterpart of condition (14). Our strategy in analyzing this problem will be as follows. We will first recall the scattering amplitudes corresponding to the two-body bulk scattering matrix (7) and the reflection coefficient (15). From these amplitudes one can recover the underlying boundary algebra \( \mathcal{B}_R \). The final step will be to check that Eq. (5) in the Fock representation \( \mathcal{F}_{R,B} \) of \( \mathcal{B}_R \) provides the solution of (1), (2), (9), (10).

The scattering theory of one-dimensional integrable systems in the presence of a reflecting boundary has been developed by Cherednik [11] and successfully applied more recently by Ghoshal and Zamolodchikov [12]. The following picture emerges from these investigations. Let \( |k_1,\ldots,k_n\rangle^{\text{in}} \) be an in-state, representing \( n \) particles coming from \( x = +\infty \) and thus having negative momenta \( k_1 < k_2 < \ldots < k_n < 0 \). These particles interact among themselves before and after being reflected by the wall at \( x = 0 \), giving rise to an out-state \( |p_1,\ldots,p_m\rangle^{\text{out}} \) composed of particles traveling toward \( x = +\infty \) and thus having positive momenta \( p_1 > p_2 > \ldots > p_m > 0 \). The transition amplitude between these states vanishes unless \( n = m \) and \( p_i = -k_i, \ i = 1,\ldots,n \). Therefore, not only the total momentum, but each momentum is separately reflected. According to [11], the scattering amplitude is
\[
\text{out} \langle p_1,\ldots,p_m|k_1,\ldots,k_n|\text{in} = \delta_{mn} \prod_{i=1}^{n} 2\pi\delta(p_i + k_i) B(p_i) \times \prod_{i,j=1, i < j} R(p_i, p_j) R(p_i, -p_j) .
\]  
The \( R \)-factors describe the interactions among the particles in the bulk, while the \( B \)-factors take into account the reflection from the wall. The crucial point now is that there exists an algebra \( \mathcal{B}_R \) encoding the scattering amplitudes (16), in perfect analogy to the case without boundary, where (13) are related to \( \mathcal{A}_R \). \( \mathcal{B}_R \) is the boundary algebra we are looking for. It plays the main role in the quantum inverse scattering approach to the boundary valued problem (1), (2) and has been introduced and investigated in a more general context in [13]. In the specific case under consideration, \( \mathcal{B}_R \) is generated by \( \{a(k),a^\ast(k),b(k)\} \). These generators satisfy quadratic exchange relations, which can be conveniently grouped in two sets. The first one is
\[
a(k_1) a(k_2) - R(k_2,k_1) a(k_2) a(k_1) = 0 ,
\]
\[
a^\ast(k_1) a^\ast(k_2) - R(k_2,k_1) a^\ast(k_2) a^\ast(k_1) = 0 ,
\]
\[
a(k_1) a^\ast(k_2) - R(k_1,k_2) a^\ast(k_2) a(k_1) = 2\pi\delta(k_1 - k_2) + b(k_1) 2\pi\delta(k_1 + k_2) .
\]  

and strongly resembles (6), except for the presence of the so called boundary generator $b(k)$ in the right hand side of the last equation. The second set of constraints describes the exchange relations of $b(k)$ and reads

$$a(k_1) b(k_2) = R(k_2, k_1) R(k_1, -k_2) b(k_2) a(k_1),$$

$$b(k_1) a^*(k_2) = R(k_2, k_1) R(k_1, -k_2) a^*(k_1) b(k_2).$$

(18)

We observe that the coupling constant $g$ enters the algebra directly through the exchange factor $R$, while there is still no reference to the boundary parameter $\eta$; it determines which of the inequivalent Fock representations of $\mathcal{B}_R$ must be chosen. Indeed, as explained in details in [13], a Fock representation $\mathcal{F}_{R,B}$ of $\mathcal{B}_R$ is characterized by a vacuum state $\Omega$, such that:

1. $\Omega$ is annihilated by $a(k)$;
2. $\Omega$ is cyclic with respect $a^*(k)$;
3. $\Omega$ is an eigenvector of $b(k)$ with eigenvalue $B(k)$.

One must distinguish the c-number reflection coefficient $B(k)$ from the boundary generator $b(k)$, which according to Eqs. (18) does not even commute with $(a(k), a^*(k))$.

Recapitulating, the mere fact that our system has a boundary shows up at the algebraic level, turning the ZF algebra $A_R$ into the boundary algebra $\mathcal{B}_R$. According to point 3 above, the details of the boundary condition (2) (namely, the value of the parameter $\eta$) enter at the representation level through the reflection coefficient $B(k)$. In the Fock space $\mathcal{F}_{R,B}$ one has [13]

$$a(k) = b(k) a(-k),$$

(19)

which is the correct quantum analogue of Eq. (14) and descends from the so called reflection automorphism of $\mathcal{B}_R$, established in [13].

The connection with Cherednik’s scattering theory is obtained through the identification

$$|k_1, \ldots, k_n \rangle_n = a^*(k_1) \ldots a^*(k_n) \Omega,$$

$$k_1 < \ldots < k_n < 0,$$

$$|p_1, \ldots, p_n \rangle_n = a^*(p_1) \ldots a^*(p_n) \Omega,$$

$$p_1 > \ldots > p_n > 0.$$

One finds in fact, that the amplitudes

$$(a^*(p_1) \ldots a^*(p_n) \Omega, a^*(k_1) \ldots a^*(k_n) \Omega),$$

where $(\cdot, \cdot)$ is the scalar product in $\mathcal{F}_{R,B}$, precisely reproduce the right hand side of Eq. (16).

The final step of our consideration is to show that inserting in Eq. (5) the generators $(a(k), a^*(k))$ in the Fock representation $\mathcal{F}_{R,B}$ of $\mathcal{B}_R$ and restricting to $x > 0$, one gets a field $\Phi(t, x)$ which solves Eqs. (1), (2) and satisfies the commutation relations (9), (10). The details in verifying the validity of this statement are given in [10]. Here we shall focus on the essential points only. First of all, we would like to fix an appropriate domain $\mathcal{D} \subset \mathcal{F}_{R,B}$ for the quantum fields. For this purpose we denote by $\mathcal{D}^n$ with $n \geq 1$ the $n$-particle subspace

$$\left\{ \int dp_1 \ldots dp_n f(p_1, \ldots, p_n) a^*(p_1) \ldots a^*(p_n) \Omega : f \in \mathcal{F}(\mathbb{R}^n) \right\}.$$

Setting $\mathcal{D}^n = \{ v \Omega : v \in \mathbb{C} \}$, we then define $\mathcal{D}$ as the linear space of sequences $\varphi = (\varphi^{(0)}, \varphi^{(1)}, \ldots, \varphi^{(n)}, \ldots)$ with $\varphi^{(n)} \in \mathcal{D}^n$ and $\varphi^{(n)} = 0$ for $n$ large enough. The last condition and the normal ordered structure of $\Phi^{(n)}(t, x)$ directly imply that the series (3) converges in mean value on $\mathcal{D}$ for any $g > 0$. By cyclicity of the vacuum, $\mathcal{D}$ is dense in $\mathcal{F}_{R,B}$ and one can show [10] that also Eqs. (1), (2) are satisfied in mean value on $\mathcal{D}$. In fact,

$$(i \partial_t + \partial_x^2)(\varphi_1, \Phi(t, x) \varphi_2) = 2 g (\varphi_1, \Phi \Phi^* \Phi : (t, x) \varphi_2),$$

$$\lim_{x \to 0} (\partial_t - \eta)(\varphi_1, \Phi(t, x) \varphi_2) = 0,$$

(20)

hold for any $\varphi_1, \varphi_2 \in \mathcal{D}$ and $x > 0$. The proof of Eq. (20) is similar in spirit to that given by Davies [4] for the NLS on $\mathbb{R}$. The novelty consists in evaluating the contributions of the boundary generator $b$ stemming from the exchange of $a$ with $a^*$. In this respect the conditions $x > 0$ and $\eta \geq 0$ show to be essential.

As it should be expected, the fields $\Phi(t, x)$ and $\Phi^*(t, x)$ have to be regarded as sesquilinear forms
on \( \mathcal{D} \). In order to deal with operators, one has to take the averages

\[ \Phi(t,h) = \int dx \Phi(t,x) h(x), \]

\[ \Phi^\ast(t,h) = \int dx \Phi^\ast(t,x) h(x), \]

with \( h \in \mathcal{D}(\mathbb{R}) \) and such that supp\( h \subset \mathbb{R}_+ \). One has by construction

\[ \Phi(t,h) \mathcal{D}^{(0)} = 0, \quad \Phi(t,h) \mathcal{D}^{(n)} \rightarrow \mathcal{D}^{(n-1)}, \quad n \geq 1, \]

\[ \Phi^\ast(t,h) \mathcal{D}^{(n)} \rightarrow \mathcal{D}^{(n+1)}, \quad n \geq 0. \]

\( \mathcal{D} \) is a common invariant domain for the operators \( \Phi(t,h) \) and \( \Phi^\ast(t,h) \), where the equal-time commutation relations

\[ [\Phi(t,h_1), \Phi(t,h_2)] = [\Phi^\ast(t,h_1), \Phi^\ast(t,h_2)] = 0, \]

\[ [\Phi(t,h_1), \Phi^\ast(t,h_2)] = \int dx h_1(x) h_2(x), \]

hold [10]. Moreover, the vacuum \( \Omega \) is cyclic also with respect to \( \Phi^\ast(t,h) \) for fixed \( t \).

Concerning the time-evolution, one can define on \( \mathcal{D} \) the essentially self-adjoint operator

\[ H = \int \frac{dp}{2\pi} \frac{p^2}{2} a^\ast(p) a(p), \]

which generates

\[ e^{iHt} a(k) e^{-iHt} = e^{-ik^2t} a(k), \]

\[ e^{iHt} b(k) e^{-iHt} = b(k). \]

Now, it is easily seen that

\[ \Phi(t,h) = e^{iHt} \Phi(0,h) e^{-iHt}, \]

which shows that \( H \) is actually the Hamiltonian of the NLS model on the half line. Notice the simple form (23) of \( H \) in terms of \( \{a(k), a^\ast(k)\} \) and the fact that \( b(k) \) does not evolve in time.

The nontrivial correlation functions of our system involve equal number of \( \Phi \) and \( \Phi^\ast \). From the structure of Eq. (5) it follows that for computing the exact 2\( n \)-point function one does not need all terms in the expansion (3), but at most the \( (n - 1) \)-th order contribution. For instance, the two-point function is given by

\[ \langle \Omega, \Phi(t_1,x_1) \Phi^\ast(t_2,x_2) \Omega \rangle = \int \frac{dp}{2\pi} e^{-i(p(t_1 - t_2))} \times \left[ e^{i(p(x_1 - x_2))} + B(p) e^{i(p(x_1 + x_2))} \right], \]

and coincides with that of the non-relativistic free field on the half line. In that context \( \eta \) parametrizes a family of self-adjoint extensions of the Laplacian on \( \mathbb{R}_+ \). The nontrivial scattering is consistently described by the 2\( n \)-point correlation functions for \( n \geq 2 \). These functions differ from the free ones and their on-shell limit leads [10] to the transition amplitudes (16), which completes the picture and concludes our quantum field theory description of the NLS model on \( \mathbb{R}_+ \).

Summarizing, we have shown above that the quantum inverse scattering transform (5) works for solving the boundary valued problem (1), (2), provided that the ZF algebra \( \mathcal{A}_B \) is replaced by the boundary algebra \( \mathcal{A}_b \). It will be interesting to extend our construction to the case \( \eta < 0 \), where the presence of boundary bound states must be taken into account.

The results of the present work confirm that boundary algebras find application not only to scattering theory [13], but also to the construction of off-shell interacting fields in integrable systems on the half line. It is worth mentioning in this respect, that our construction can be easily adapted to the complex modified Korteweg-de Vries equation. Other applications are currently under investigation.

References


On the decay law for unstable open systems

F. Benatti \textsuperscript{a}, R. Floreanini \textsuperscript{b}

\textsuperscript{a} Dipartimento di Fisica Teorica, Università di Trieste, Strada Costiera 11, 34014 Trieste, Italy

and Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, Italy

\textsuperscript{b} Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, Dipartimento di Fisica Teorica, Università di Trieste, Strada Costiera 11, 34014 Trieste, Italy

Received 25 February 1998

Editor: R. Gatto

Abstract

We use (nonconservative) dynamical semigroups to investigate the decay law of a quantum unstable system weakly coupled with a large environment. We find that the deviations from the classical exponential law are small and can be safely ignored in any actual experiment. © 1998 Published by Elsevier Science B.V. All rights reserved.

The decay law of an unstable microscopic system can be rather well described by an exponential function; this result can be easily justified on the basis of classical probabilistic considerations. However, microscopic systems should be described by quantum mechanics and it is well known that in quantum theory the exponential decay law can not be valid for all times; in particular, it surely fails for very short and very long times. Indeed, rather general considerations assure that the quantum decay law can be described by a three-step function: a Gaussian law at short times, the classical exponential law at intermediate times and finally a power law at longer times (e.g. see [1–3] and references therein).

Although various experiments have been devised in order to obtain evidence for discrepancies with the exponential law, none have been so far actually detected [4]. In this respect, a system that has attracted a lot of interest both theoretically and experimentally is the neutral kaon system. The $K^0, \bar{K}^0$ system has proven to be one of the most fruitful systems for testing fundamental symmetries, like $CP$ and $CPT$. In parametrizing violations of these symmetries, one usually takes for granted the exponential decay law and uses an effective theory to describe the kaon system [5].

In the following we shall examine to what extent the exponential decay law can be trusted in actual experiments. More precisely, we shall study the deviations from the exponential law for small and large times, taking also into account possible effects due to incoherent interactions with the environment.

The general idea that is at the basis of our considerations is that unstable systems can be viewed as specific examples of open quantum systems. These systems can be modeled in general as being small subsystems in weak interaction with large environments. Although the global time evolution of the closed compound system is described by an unitary transformation, the reduced dynamics of the subsystem, obtained by the elimination of the environment degrees of freedom, usually develops some sort of
dissipation and irreversibility. Under mild assumptions, the reduced evolution is realized by oneparameter (= time) maps acting on the states of the system, conveniently described by density matrices, with forward in time composition (semigroup property) and the additional characteristic of being completely positive. This set of transformations forms a so-called dynamical semigroup [6–10].

This rather universal and general formalism has been recently adopted to treat effective dynamics for the kaon system that transform pure states into mixed ones [11–16]. The physical motivations behind such an approach are based on quantum gravity, that predicts loss of quantum coherence at the Planck’s scale [17]. These generalized time-evolutions lead to CP and CPT violating effects that could be in the reach of the next generation of neutral kaon experiments [15,16].

In these treatments, the environment was assumed not to contribute to the decay of the kaons, that was effectively described by the standard exponential law. However, possible effects of the environment on the decay process itself are surely conceivable. As we shall see, they can be studied using again the formalism of dynamical semigroups. It will turn out that these environment effects do not modify the short-time behaviour of the decay law, but only the exponential and power law regimes.

However, in the realistic hypothesis of a weak coupling between unstable system and environment, these modifications are tiny or occur for too large times for any practical considerations. Therefore, the exponential decay law captures rather well the essential features of the decay process of an unstable system and can be certainly used with confidence in any actual experiment involving weak-decaying particles, like the neutral K mesons.

For our considerations, we choose to describe the states of a quantum system evolving in time by means of density matrices $\rho(t)$. Given an initial state $\rho$ and a time-independent hermitian hamiltonian $H$, the standard time evolution is described by

$$\rho(t) = e^{-iHt} \rho e^{iHt},$$

solution of the Liouville-von Neumann equation

$$\frac{\partial \rho(t)}{\partial t} = -i[H, \rho(t)].$$

In the case of an unstable system, it is custom to split $H$ as

$$H = H^{(0)} + H^{(1)},$$

where $H^{(0)}$ is the unperturbed hamiltonian, while $H^{(1)}$ is the interaction hamiltonian that drives the decay process; the system would be stable if $H^{(1)} = 0$. In the case of the neutral K-mesons, $H^{(0)}$ can be identified with the hamiltonian of the strong interactions and $H^{(1)}$ with that of the weak interactions (possible “superweak” mixing terms should be included in $H^{(0)}$). Further, we shall call $P_u$ the projector operator on the subspace of the undecayed states. We shall also use the orthogonal projector $P_d = 1 - P_u$; it describes the transition to the space of the decayed states.

Assuming that at the beginning our unstable system is in the undecayed state $\rho = P_u \rho P_u$, the probability of finding it undecayed at time $t$ is given by

$$\mathcal{P}(t) = \text{Tr}[\rho(t) P_u].$$

In the case of the decay of a single particle, with evolution as in (1), the properties of $\mathcal{P}(t)$ have been widely studied in the literature. Here we shall extend those treatments by considering also the interaction of the decay system (not necessarily one-dimensional) with the environment. Together with the effects of the hamiltonian $H^{(1)}$, also this interaction could in principle contribute to the decay process.

As mentioned in the introductory remarks, in order to take into account these extra effects we shall treat the unstable system as an open quantum system $S$ [6–10]. As for any open system, $S$ can be thought of as interacting with a suitable environment $E$, so that the global system $S + E$ is closed. This evolves in time according to a group of unitary operators as in (1), governed by a total hamiltonian $H_{S + E}$, which is the sum of the hamiltonian $H_S$ of the subsystem, of the hamiltonian $H_E$ of the environment and of the interaction hamiltonian $H_{SE}$ between them. A reduced dynamics for the subsystem $S$ can be consistently obtained when $S$ and $E$ are assumed to be uncorrelated at the moment of the formation of the unstable system and therefore the state of the total system is simply: $\rho_{S + E} = \rho \otimes \rho_E$. 


In such cases, by tracing over the the environment degrees of freedom one gets linear, completely positive maps $\rho \rightarrow \tilde{\gamma}_t[\rho]$ on the states of the subsystem, where

$$\tilde{\gamma}_t[\rho] \equiv \text{Tr}_E[e^{-iH_t\tau}(\rho \otimes \rho_r)e^{iH_t\tau}], \quad (5)$$

These transformations do not have any simple composition law and they usually contain memory effects. However, when the interaction with the environment is weak, they reduce to completely positive, dissipative dynamical maps $\gamma_t: \rho \rightarrow \rho(t)$, which obey a semigroup composition law, $\gamma_t \circ \gamma_{t_2} = \gamma_{t+t_2}$, $t_1,t_2 \geq 0$, and satisfy the condition of entropy increase: $dS/dt \geq 0$, $S(t) = -\text{Tr}[\rho(t)\ln\rho(t)]$. Moreover, these semigroups are generated by equations of a very specific type that can be explicitly given:

$$\frac{d\rho(t)}{dt} = R\rho(t) + (\rho(t)R^\dagger + \sum_k A_k\rho(t)A_k^\dagger), \quad (6)$$

where the set of (bounded) operators $R$ and $A_k$ are such that: $R + R^\dagger + \sum_k A_k^\dagger A_k \leq 0$. They are probability preserving or not depending on whether this combination is identically zero or not; they are called quantum dynamical semigroups. Notice that this description of open systems is rather general and the effective subdynamics is essentially independent of the type of the environment.

In the case of unstable high energy particle systems, a natural choice for $E$ could be the gravitational field, whose effects are usually neglected in the theory of elementary particles because of the smallness of its coupling. Its quantum fluctuations at Planck's length could nevertheless act as a weak coupled environment producing detectable effects in elementary particle interactions [17,11]. It is also worth mentioning that the effects of such fluctuations, realized via the space-time foam, can be effectively described with an heat bath, the most natural of all environments [18].

In view of all above considerations, we shall now generalize the standard quantum mechanical evolution equations for an unstable system by adding to (2) a linear piece of the form given by the r.h.s. of (6). Since this additional piece should describe the contribution of the environment to the decay process, it must be proportional to $\rho(t) - P_u\rho(t)P_u$. In fact, this term would project $\rho(t)$ out of the space of undecayed states. Thus, we shall study the generalized time evolution described by the equation

$$\frac{d\rho(t)}{dt} = -i[H,\rho(t)] - \lambda(\rho(t) - P_u\rho(t)P_u), \quad (7)$$

with $\lambda$ a positive “coupling” constant and $H$ as in (3). We shall consider weak coupled environments, and therefore assume $\lambda$ to be much smaller than any typical energy scale in the hamiltonian. In the case of the $K^0 - \bar{K}^0$ system, assuming the dissipative term in (7) of gravitational origin, dimensional arguments suggest $\lambda$ to be at most of order $m_K^2/m_P$, where $m_K$ is the kaon rest mass and $m_P$ is Planck's mass. The evolution Eq. (7) describes nonconservative dynamical maps [6] Indeed, probability is not conserved, $(d/dt)\text{Tr}[\rho(t)] = -\lambda\text{Tr}[\rho(t)P_u] \leq 0$, as it should be for an unstable system; however, this violation is small, since $\lambda$ is small.

Equations of the form (7) were used before in discussing one-dimensional unstable systems [1]. The motivation for the introduction of the nonstandard term was there attributed to the interaction of the decaying system with the measuring apparatus. While this point of view is surely viable, we stress that our interpretation of (7) as describing the evolution of an unstable system is completely phenomenological in nature. In particular, we do not make any specific assumption on the phenomena responsible for the appearance of the second term in the r.h.s. of (7).

In order to proceed in the study of Eq. (7), it is convenient to introduce a vector notation, rewriting the density matrix $\rho$ describing the state of the unstable system as the vector $|\rho\rangle$. To any operator acting on the state $\rho$, one can define a corresponding operator acting on the vector $|\rho\rangle$. In particular, one can define the projector $\Pi_u$,

$$\Pi_u|\rho\rangle \equiv |P_u\rho P_u\rangle, \quad \Pi_u^2 = \Pi_u, \quad (8)$$

and its orthogonal complement: $\Pi_d = 1 - \Pi_u$. By introducing the Liouville operator $L_H$ corresponding to the hamiltonian $H$,

$$L_H|\rho\rangle \equiv |[H,\rho]\rangle, \quad (9)$$
one can rewrite (7) as a Schrödinger like equation:

$$\frac{d}{dt} |\rho(t)\rangle = L |\rho(t)\rangle,$$

(10)

with

$$L \equiv H - i\lambda \Pi_d. $$

(11)

Although the statements and the conclusions obtained below using this new formalism can be rigorously justified, for sake of simplicity we shall keep mathematical considerations to a minimum. As we shall see in the following, this formalism results particularly appropriate in the study of unstable systems for which the space of the undecayed states is not one-dimensional, as in the case of the neutral kaons.

We are ultimately interested in describing the properties of the probability $P(t)$ in (4). Therefore, one should concentrate on the study of the time evolution of the projected vector:

$$|\rho(t)\rangle_a = \Pi_a |\rho(t)\rangle.$$

(12)

Using the Laplace transformed vector $|\tilde{\rho}(s)\rangle$, the corresponding Lippmann-Schwinger equation reads:

$$s + i \left( L_{uu} + L_{ud} \right) + \frac{1}{i s - L_{dd} + i\lambda \Pi_d} L_{du} |\tilde{\rho}(s)\rangle_a = |\tilde{\rho}(0)\rangle_a,$$

(13)

where $s$ is the Laplace variable and

$$L_{uu} = \Pi_a L_H \Pi_a, \quad L_{ud} = \Pi_a L_H \Pi_d,$$

$$L_{du} = \Pi_d L_H \Pi_a, \quad L_{dd} = \Pi_d L_H \Pi_d.$$  

(14)

The action of these operators on any state $|\rho\rangle$ is well-defined. For instance, using the shorthand notation $\Theta_{ij} = P_i \Theta P_j, i, j = u, d$, with $\Theta$ a generic operator, one finds:

$$L_{ud} |\rho\rangle = |H_{ud} \rho_{ud} - \rho_{ud} H_{du}\rangle,$$

$$L_{du} |\rho\rangle = |H_{du} \rho_{du} - \rho_{du} H_{ud}\rangle. $$

(15a)

(15b)

According to our hypothesis, the decay process is driven by the interaction hamiltonian $H^{(1)}$ and the coupling to the environment. In particular, the decaying system would be actually stable for $H^{(1)} = 0$ and $\lambda = 0$. This immediately implies that $H_{ud} \equiv H_{ud}^{(1)}$, $H_{uu} \equiv H_{uu}^{(1)}$ and $H_{du} \equiv H_{du}^{(0)}$. Furthermore, as $H^{(1)}$ is supposed to be small (in the case of the neutral kaons, $H^{(1)}$ is indeed the weak hamiltonian), we shall study approximated solutions of (13), taking into account only terms up to second order in $H^{(1)}$; this is also the approximation that is usually adopted in the description of decaying particles in standard quantum mechanics.

At this point, in order to simplify the formulas, we also assume that the space of the undecayed states is degenerate in energy: $H_{uu} \equiv H_{uu}^{(1)} = E_0 \Pi_u$. In the case of the neutral kaons, this is not really a restriction since $K^0$ and $\bar{K}^0$ have the same rest mass. Then, by acting on a generic state $|\rho\rangle_a$, one can prove that:

$$L_{uu} + L_{ud} \frac{1}{is + \lambda \Pi_d - L_{dd}} L_{du}$$

$$= \mathcal{L}_{W(s)} + O((H^{(1)}))^3,$$

(16)

where $W(s)$ is the effective non-hermitian "hamiltonian"

$$W(s) = H_{ud}^{(1)} \frac{1}{is + \lambda + E_0 - H_{uu}^{(1)}} H_{dd}^{(1)},$$

(17)

and $\mathcal{L}_{W(s)}$ is the corresponding generalized Liouville operator

$$\mathcal{L}_{W(s)} |\rho\rangle = [W(s) \rho - \rho W^{\dagger}(s)].$$

(18)

This is not surprising since, in view of (15), the generalized operators $L_{ud}$ and $L_{du}$ themselves are of order $H^{(1)}$. Thus, up to second order terms in $H^{(1)}$, Eq. (13) becomes

$$(s + i \mathcal{L}_{W(s)}) |\tilde{\rho}(s)\rangle_a = |\tilde{\rho}(0)\rangle_a.$$  

(19)

Using the inverse Laplace transform, one can then write

$$|\rho(t)\rangle_a = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} ds e^{st} \left[ \frac{1}{s + i \mathcal{L}_{W(s)}} \right] |\tilde{\rho}(0)\rangle_a,$$

(20)

where $c$ must be chosen so that the integration path in the complex $s$-plane lies to the right of all singularities of the generalized operator $[s + i \mathcal{L}_{W(s)}]^{-1}$.

Before proceeding further, let us first deduce from (20) the standard exponential decay law. The singularities of the integrand in (20) are related to those of the Liouville operator $\mathcal{L}_{W(s)}$, and therefore to those of $W(s)$ in (17). Indeed, using the definition (18),
one can deduce that the spectrum of $\mathcal{L}_{W(s)}$ coincides with the difference of the spectra of $W(s)$ and $W'(s)$. The operator $W(s)$ is analytic in the complex $s$-plane except when the denominator in (17) vanishes. Since typically the spectrum of $H^{(0)}_{dd}$ is continuum, the singularities of $W(s)$ lay on the imaginary axis, where usually there is a cut. Therefore, one can conclude that also $[s + i\mathcal{L}_{W(s)}]^{-1}$ has generally a cut on the imaginary axis of the $s$-plane. One can then move the integration path in (20) almost to coincide with this axis.

Further, one notice that when $H^{(1)} = 0$ and $\lambda = 0$ the generalized Liouville operator $\mathcal{L}_{W(s)}$ becomes the null operator, so that only a pole at the origin occurs in the integrand of (20). In this case, we obtain: $|\rho(t)\rangle_a = |\rho(0)\rangle_a$, as it should be, since now the system is stable. In presence of interactions, this pole moves into the second sheet, but remains close to the imaginary axis for small $H^{(1)}$ and $\lambda$ [3]. One then expects that this pole continues to give the main contribution to the integral. This is the so called Weisskopf-Wigner approximation, that gives rise to the exponential decay law for all times. Indeed, within this approximation and using (18), the integral in (20) gives

$$|\rho(t)\rangle_a = e^{-i\mathcal{L}_{W(s)}} |\rho(0)\rangle_a,$$

or equivalently

$$\rho_a(t) = e^{-iH^{(1)}} \rho_a(0) e^{iH^{(1)}},$$

where the effective hamiltonian $H_W$ takes the form:

$$H_W = H^{(0)}_{uu} + H^{(1)}_{dd} \frac{1}{E_0 - H^{(0)}_{dd} + i\lambda H^{(1)}_{dd}}.$$  \hfill (23)

This hamiltonian is not hermitian and can be written as $H_W = M - i\Gamma/2$, with $M$ and $\Gamma$ hermitian and positive. Indeed, by putting the system in a finite box so that the spectrum of $H^{(0)}_{dd}$ becomes discrete, one explicitly finds:

$$[M]_{\alpha\beta} = E_0 \delta_{\alpha\beta}$$

$$+ \sum_k \langle \alpha | H^{(1)} | k \rangle \left( \frac{E_0 - E_k}{(E_0 - E_k)^2 + \lambda^2} \right) \times \langle k | H^{(1)} | \beta \rangle,$$  \hfill (24a)

$$\left[ \Gamma \right]_{\alpha\beta} = 2 \sum_k \langle \alpha | H^{(1)} | k \rangle \left( \frac{\lambda}{(E_0 - E_k)^2 + \lambda^2} \right) \times \langle k | H^{(1)} | \beta \rangle,$$  \hfill (24b)

where we have used greek (latin) indices to label undecayed (decay-products) states. The entries of the matrix

$$[\sigma(E)]_{\alpha\beta} = \sum_{E_\in E} \langle \alpha | H^{(1)} | k \rangle \langle k | H^{(1)} | \beta \rangle,$$  \hfill (25)

are usually found to be piecewise differentiable in the variable $E$ when the volume of the box becomes infinite [2]. In this limit, by setting

$$[\omega(E)]_{\alpha\beta} = \frac{d[\sigma(E)]_{\alpha\beta}}{dE},$$  \hfill (26)

the sums in (24) can be substituted by integrals:

$$[M]_{\alpha\beta} = E_0 \delta_{\alpha\beta}$$

$$+ \int_{E_m}^{E_\infty} dE [\omega(E)]_{\alpha\beta} \left( \frac{E_0 - E}{(E_0 - E)^2 + \lambda^2} \right),$$  \hfill (27a)

$$\left[ \Gamma \right]_{\alpha\beta} = 2 \int_{E_m}^{E_\infty} dE [\omega(E)]_{\alpha\beta} \left( \frac{\lambda}{(E_0 - E)^2 + \lambda^2} \right),$$  \hfill (27b)

where $E_m$ is the lowest eigenvalue of $H^{(0)}_{dd}$. For $\lambda$ small, using

$$\frac{x}{x^2 + \lambda^2} = P \frac{1}{x} + \lambda \pi \delta'(x) + O(\lambda^3),$$  \hfill (28a)

$$\frac{x}{x^2 + \lambda^2} = \pi \delta(x) + \lambda P \frac{1}{x^2} + O(\lambda^2),$$  \hfill (28b)

with $P$ indicating principal value, the Eqs. (27) finally become:

$$[M]_{\alpha\beta} = \left[ \mu(E_0) \right]_{\alpha\beta} + \lambda \pi \left[ \omega'(E_0) \right]_{\alpha\beta},$$  \hfill (29a)

$$\left[ \Gamma \right]_{\alpha\beta} = 2 \pi \left[ \omega(E_0) \right]_{\alpha\beta}$$

$$+ 2 \lambda \left( \delta_{\alpha\beta} - \left[ \mu'(E_0) \right]_{\alpha\beta} \right),$$  \hfill (29b)
where the dash signifies derivative with respect to $E_\alpha$ and
\[
\left[ \mu(E_\alpha) \right]_{\alpha\beta} = E_\alpha \delta_{\alpha\beta} + \int_{E_\alpha}^{\infty} dE \left\{ \omega(E) \right\}_{\alpha\beta} P \frac{1}{E_0 - E}. \tag{30}
\]

When $\lambda = 0$, $M$ and $\Gamma$ reduce to their standard Wiesskopf-Wigner expressions. The effect of the environment is to modify these expressions by adding terms that are in principle calculable using field theory techniques [19–21]. The actual evaluation of the various terms in (29) requires the adoption of a specific microscopic model for the interacting hamiltonian and is certainly beyond the purpose of the present work.

As stressed at the beginning, the exponential decay law can not hold for all times. We shall now go back to the Eq. (19), which is exact up to second order terms in $H^{(1)}$, and try to evaluate the corrections to the Weisskopf-Wigner approximation. To this purpose, let us add and subtract to the l.h.s. of (19) the term $\mathcal{L}_{H_0}$ [22]. Within our approximation, one can then write:
\[
\left[ s + i \mathcal{L}_{H_0} \right]^{-1} = \left[ s + i \mathcal{L}_{H_0} \right]^{-1} - \frac{i}{s^2} \left( \mathcal{L}_{H_0} - \mathcal{L}_{H_0} \right). \tag{31}
\]

By acting on a generic state $|\rho\rangle$$_u$, using simple algebra one can prove that:
\[
-\frac{i}{s^2} \left( \mathcal{L}_{H_0} - \mathcal{L}_{H_0} \right) |\rho\rangle_u = |V(s)\rho_{uu} - \rho_{uu} V^\dagger(s)\rangle, \tag{32}
\]
where
\[
V(s) = H^{(1)}_{uu} \frac{1}{(E_0 - H^{(0)}_{dd} + i\lambda)^2} \times \left[ \frac{1}{i(s + \lambda) + E_0 - H^{(0)}_{dd}} - \frac{1}{is} \right] H^{(1)}_{du}. \tag{33}
\]

Inserting these results in (20) and performing the $s$-integration, one obtains an effective evolution for the projected density matrix $\rho_{uu}$ of the form
\[
\rho_{uu}(t) = U(t) \rho_{uu}(0) U^\dagger(t), \tag{34}
\]

where
\[
U(t) = e^{-i H_{uu} t} + H^{(1)}_{uu} \frac{1}{(E_0 - H^{(0)}_{dd} + i\lambda)^2} \times \left[ e^{-i(H^{(0)}_{dd} - i\lambda)t} - e^{-iE_0 t} \right] H^{(1)}_{du}. \tag{35}
\]

This evolution satisfies the correct boundary condition, $U(0) = P_u$; it contains contributions up to second order in $H^{(1)}$ and to all orders in $\lambda$. It is dominated by the exponential Wiesskopf-Wigner term; however, the additional pieces are relevant for small and large times.

By expanding (35) for small $t$, one gets:
\[
U(t) \sim (1 - itE_0) P_u - \frac{t^2}{2} \left( E_0^2 P_u + H^{(1)}_{uu} H^{(1)}_{du} \right). \tag{36}
\]

Inserting this result in (34), the probability (4) takes the form:
\[
\mathcal{P}(t) = 1 - \left( t/\tau_G \right)^2 \equiv 1 - \frac{t^2}{2} \text{Tr} \left[ H^{(1)}_{uu} H^{(1)}_{du} \rho_{uu}(0) \right], \tag{37}
\]

and therefore has a Gaussian behaviour; the Gaussian width $\tau_G^{-2}$, that can be rewritten as $\text{Tr} \left[ H_{uu} H_{du} \rho_{uu}(0) \right]$, gives the spread in energy $(\Delta E)^2$ of the initial state $\rho_{uu}(0)$. Furthermore, $\mathcal{P}(t)$ is independent from $\lambda$, so that the small-time decay law is unaffected by the interaction with the environment. This result is physically understandable: the time is too short to allow the environment to play a role in the decay, which, in this early stages, is totally driven by $H^{(1)}$. This is also in agreement with our starting assumption that the environment does not disturb the preparation of the decaying system.

The times for which the Gaussian behaviour of $\mathcal{P}(t)$ can be, at least in principle, experimentally detected are in general very small. A comparison with the exponential behaviour in (22), which gives
\[
\mathcal{P}(t) = 1 - t/\tau \equiv 1 - t \text{Tr} \left[ \Gamma \rho_{uu}(0) \right], \tag{38}
\]

indicates that this could happen only for times smaller than $t_m = \tau_G^2/\tau = 1/(\Delta E)^2 \tau$. However, by taking into account the Heisenberg time-energy uncertainty principle, $\Delta E \Delta t \geq 1$, one finds $t_m$ to be actually
smaller than the time interval $\Delta t$ necessary to complete any measurement [1,23]. Therefore, it is practically impossible to detect deviations from the standard exponential decay law at small times in any actual experiment involving elementary particles (nevertheless, in suitable atomic systems the situation might be different, see [24,25]).

In studying the large time behaviour of $U(t)$, it is convenient to rewrite (35) as:

$$U(t) = \left[1 - R(0)\right] e^{-iH_0 t} + R(t),$$

where

$$R(t) = H^{(1)}_{ad} \frac{1}{(E_0 - H^{(0)}_{ad} + i\lambda)^2} e^{-i(H^{(0)}_{ad} - i\lambda)t} H^{(1)}_{ad}. \quad (40)$$

For large times, the prefactor to the Wisskopf-Wigner term, which is actually equal to $dH_0(E_0)/dE_0$, can be reabsorbed in a normalization of the states. By putting the system in a box and using again the definitions (25) and (26), in the infinite-volume limit the correction term $R(t)$ takes the form

$$R(t) = \int_{E_m}^{\infty} dE \omega(E) \frac{e^{-i(E - i\lambda)t}}{(E_0 - E + i\lambda)^2}. \quad (41)$$

Due to phase space limitations, the behaviour on threshold of $\omega(E)$ can be usually approximated by:

$$\omega(E) \approx \omega(E_0) \left(\frac{E}{E_m} - 1\right)^\delta, \quad \delta > 0. \quad (42)$$

Inserting this expression in (41), together with a convergent factor $e^{-\epsilon E}$, $\epsilon \ll t$, the integral in $R(t)$ can be evaluated exactly in terms of Whittaker functions [26]. Taking the large $t$ limit, one finally obtains

$$R(t) \sim \Gamma(\delta + 1) \frac{E_m \omega(E_0)}{(E_0 - E_m + i\lambda)^2} \frac{e^{-i(E_m - i\lambda)t}}{(itE_m)^{\delta + 1}}. \quad (43)$$

Therefore, for large times the effective evolution operator $U(t)$ exhibits a power law behaviour, modulated by an exponential [22]. Notice that the probability $P(t)$ in (4) has a more complicated behaviour due to the interference effects between the two terms in (39). In fact, besides the standard Wisskopf-Wigner term proportional to $e^{-1/1}$ and the power-like term proportional to $t^{-2(\delta + 1)}$, $R(t)$ contains also an oscillating term that, taking for simplicity $\lambda = 0$, is proportional to $e^{-i/1} + i(\lambda + 1)$. In another context, this intermediate behaviour has also been noticed in [24].

In order to estimate the region in which the long-time correction (43) supersedes the standard exponential decay term, one has to consider the magnitude of the correction, given by the modulus $|R(t)|$ of the operator in (43). Recalling (29b), one has

$$|R(t)| = \left|\Gamma - 2\lambda \Delta \Gamma\right| \times \frac{E_m}{(E_0 - E_m)^2 + \lambda^2 (iE_m)^{\delta + 1}}. \quad (44)$$

where $\Delta \Gamma = 1 - \mu(E_0)$. This result should be compared with the standard exponential term $e^{-\Gamma t/2}$. In order to do this, let us choose a basis in which $\Gamma$(or better the spectral operator $\alpha$) is diagonal, and label with the index $\nu$ the corresponding eigenvalues. In the case of the neutral kaon system, $\alpha$ takes the two values $S$ and $L$, which refer to the $K_S$ and $K_L$ states. Assuming the difference $E_0 - E_m$ of the same order of $E_m$, one finds that the power law behaviour dominates for times larger than $\tau_\nu/\Gamma_\nu$, where $\tau_\nu$ is implicitly given by the following equation

$$\tau_\nu = 2 \left(1 + 2\lambda \Gamma_\nu\right) (\delta + 1)\ln \tau_\nu + 2 \left(1 + 2\lambda \Gamma_\nu\right) (\delta + 2)\ln \frac{E_m}{\Gamma_\nu} + 2\lambda \Delta \Gamma_\nu \Gamma_\nu. \quad (45)$$

Since $E_m/\Gamma_\nu$ is usually very large ($\sim 10^{15}$) in elementary particle decays, it turns out that deviations from the exponential law can be seen only after hundreds of life-times (for $\delta = 1$ and $\lambda = 0$), a region which is clearly unavailable to the experiment. Furthermore, notice that the role of the environment tends to worsen the situation, pushing the limit towards even longer times.

We can safely conclude that the role played by the environment in the decay process of an unstable system is marginal and can not be detected in actual
experiments. This conclusion results from a careful study of the starting evolution Eq. (7). Although phenomenological in nature, this equation encodes in a rather general and universal way possible dissipative effects due to a weak interaction with an environment. From this point of view, the description of the neutral kaon system in terms of completely positive dynamical semigroups given in [13–16] is appropriate. In particular, the deviations from the predictions of the standard Weisskopf-Wigner theory discussed there, if experimentally detected, could be really the sign of new physics.

We thank N. Paver for many illuminating discussions.

References

[18] L.J. Garay, Spacetime foam as quantum thermal bath, gr-qc/9801024.
Truncation effects in Monte Carlo renormalization group improved lattice actions

Tetsuya Takaishi a, Philippe de Forcrand b

a Hiroshima University of Economics, Hiroshima 731-01, Japan
b Swiss Center for Scientific Computing, ETH-Zürich, CH-8092 Zürich, Switzerland

Received 6 March 1998
Editor: R. Gatto

Abstract

We study truncation effects in the SU(3) gauge actions obtained by the Monte Carlo renormalization group method. By measuring the heavy quark potential we find that the truncation effects in the actions coarsen the lattice by 40–50% from the original blocked lattice. On the other hand, we find that rotational symmetry of the heavy quark potentials is well recovered on such coarse lattices, which may indicate that rotational symmetry breaking terms are easily cancelled out by adding a short distance operator. We also discuss the possibility of reducing truncation effects. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

Recently a lot of attention has been devoted to improvements of lattice discretized actions. There exist two approaches to improving actions. One is the perturbative improvement program suggested by Symanzik [1] and the other the renormalization group improvement program by Wilson [2]. Early attempts at the perturbative improvement program did not appear practical for Monte Carlo simulations [3]. Recently, however, it has been revitalized with the help of tadpole improvement [4] and actively been investigated by Monte Carlo simulations.

The renormalization group improvement program is very attractive since it can, in principle, give us lattice-artifact-free actions. In practice, however, it is not an easy task to obtain such actions since we do not know a practical way to determine them. Recent successful attempts [5] approximate the perfect action, which is defined on the renormalized trajectory, with the fixed point action (classically perfect action) obtained in the limit of $\beta \to \infty$. An advantage of this method is that the fixed point action is rather easily obtainable. Although the fixed point action is an approximation to the perfect action, this approximation turns out to be rather good in Monte Carlo tests.

Direct attempts [6,8,10,11] to obtain actions on the renormalized trajectory have been also made, using the Monte Carlo renormalization group (MCRG) method [7]. The MCRG method is useful not only for determining improved actions but also for other purposes, like monitoring the flow of couplings under a scale transforma-
tion. In lattice QCD for instance, the MCRG method has been successfully used for determining the coupling shift $\Delta \beta$ of the $\beta$ function to reveal the scaling behavior of the theory \footnote{1}. The MCRG approach uses the fact that the blocked trajectory generated by the blocking transformation reaches the renormalized trajectory after sufficiently many blocking steps and then runs along with the renormalized trajectory. Therefore the configurations generated by successive blocking transformations will correspond to an action located nearer and nearer the renormalized trajectory. Such configurations should have less artifacts than non-blocked configurations since from Wilson’s renormalization group argument we expect that continuum physics is realized on the renormalized trajectory. Actually it has been shown that rotational symmetry of the heavy quark potential is well recovered on blocked configurations \footnote{8}. If we can determine the action representing the blocked configurations, we can directly generate the configurations without blocking. The determination of the action was tackled by the canonical demon method \footnote{10,11} which produces as output a set of coupling constants. In the canonical demon method, like other determination methods, the action to be determined must be truncated to a certain local form, which may cause some error. We call the effects of this error truncation effects. Unless the truncated form of the action is sufficiently close to the real one, truncation effects may affect long-distance, physical properties of the action. In Ref \footnote{8} several SU(3) gauge actions corresponding to blocked configurations (generated by a blocking transformation) were obtained in multi-dimensional coupling space. However it is not known whether these actions still preserve the improvements of the blocked configurations or not, since these improvements may have been ruined by truncation effects. This point can only be examined by Monte Carlo simulations. In this letter, we perform Monte Carlo simulations with the actions obtained in Ref \footnote{8}, which we call MCRG improved actions, examine the rotational symmetry of the heavy quark potential, and estimate the truncation effects which appear in a physical quantity (the string tension).

2. Monte Carlo renormalization group improved actions

We briefly sketch the method used to determine the MCRG actions in Ref \footnote{8}. First, we block configurations generated with the standard Wilson action on $32^3 \times 64$ lattices. The blocking scheme employed is Swendsen’s scale factor 2 blocking scheme. This blocking scheme was optimized by multiplication by a Gaussian random SU(3) matrix so that the blocked trajectory converges to the renormalized trajectory quickly \footnote{9,8}. Actually the optimal width of the Gaussian distribution at $\beta \sim 6.0$ turned out to be 0, meaning that the Gaussian random SU(3) matrix becomes unity. This optimal width was used in the blocking. The blocking was performed twice, starting from $32^3 \times 64$ lattices at two $\beta$ values, 6.0 and 6.3, thus resulting in two sets of $8^3 \times 16$ blocked configurations.

Next, in order to determine coupling constants we apply the canonical demon method \footnote{10,11} on the blocked configurations. The canonical demon method introduces several degrees of freedom, so-called demons, which are associated with each of the coupling constants to be determined. The action $S$ can be written as

$$S = \sum_i \beta_i \bar{S}(U)$$

where $\beta_i$ is a coupling constant and $\bar{S}(U)$ is some operator consisting of Wilson loops. Hereafter for simplicity, let us assume that the action is characterized by only one operator or by one coupling constant:

$$S = \beta \bar{S}(U).$$

The demon is updated by a microcanonical simulation in the joint system, i.e. the demon and the links of a blocked configuration. In the microcanonical simulation, the total energy of the joint system, i.e. $\bar{S}(U)$ plus the demon energy $E_d$, is kept constant.

\footnote{1} See, for example, recent results of [9].
Table 1

Values of coupling constants, $\beta_{11}$ and $\beta_{12}$, used in the study. We label these actions by A, M and B as in the table.

<table>
<thead>
<tr>
<th>Action</th>
<th>$\beta_{11}$</th>
<th>$\beta_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6.1564</td>
<td>-0.6241</td>
</tr>
<tr>
<td>M</td>
<td>7.0712</td>
<td>-0.7705</td>
</tr>
<tr>
<td>B</td>
<td>7.986</td>
<td>-0.9169</td>
</tr>
</tbody>
</table>

In the canonical demon method, in order to avoid a possible finite volume error, a set of well uncorrelated blocked configurations is prepared. At a certain stage of the microcanonical simulation, we move to a new blocked configuration chosen from the set and the configuration used before is discarded. The demon is also moved to the new configuration keeping the value of the demon energy at the last update.

The probability distribution of the demon energy $P(E_d)$ in the simulation is expected to be the Boltzmann distribution:

$$P(E_d) \sim \exp(-\beta E_d)$$

where $\beta$ is the coupling constant to be determined. Using Eq. (3), we write

$$\langle E_d \rangle = \int_{E_{\min}}^{E_{\max}} E_d \exp(-\beta E_d) dE_d / Z$$

where the demon energy $E_d$ is restricted in the region, $E_{\min} < E_d < E_{\max}$, and $Z$ is the partition function. If we take $E_{\max} = -E_{\min} = E_c$, where $E_c$ stands for some constant value which we fix in the simulation, Eq. (4) will be

$$\langle E_d \rangle = 1 / \beta - E_c / \tanh(\beta E_c).$$

Finally, substituting the average value $\langle E_d \rangle$ obtained from the simulation, we determine the value of the coupling constant $\beta$ by solving Eq. (5) numerically. The extension to the multi-coupling form of Eq. (1) is straightforward.

Fig. 1. Location of the actions in the $\beta_{11} - \beta_{12}$ plane. Points A and B were obtained in [8]. Point M is interpolated at the mid-point between A and B. Symanzik and Iwasaki actions are also indicated by solid lines in the figure.
In our study we take MCRG actions obtained in two-dimensional coupling space. The actions are written as
\[ S = \text{Re}\left( \beta_{11} \sum \text{Tr}(1 \times 1 \text{ Wilson loop})/3 + \beta_{12} \sum \text{Tr}(1 \times 2 \text{ Wilson loop})/3 \right). \]  

The values of the couplings \( \beta_{11} \) and \( \beta_{12} \) are listed in Table 1. The actions \( \text{A} \) and \( \text{B} \) come from Ref [8], which are obtained from the configurations blocked twice at \( \beta = 6.00 \) and \( \beta = 6.30 \) respectively. We also use an interpolated action \( \text{M} \) located half-way between actions \( \text{A} \) and \( \text{B} \). Fig. 1 shows the locations of these actions in the \( \beta_{11} - \beta_{12} \) plane. The ratio of the two couplings, \( |\beta_{12}/\beta_{11}| \), is bigger than that of the Symanzik tree-level improved actions \( \text{12} \) and Iwasaki action \( \text{13} \).

### 3. Heavy quark potentials

We employ \( 8^3 \times 16 \) lattices which is the same lattice size with the original blocked lattices. We calculate static potentials between a heavy quark and antiquark pair from the exponential fall-off of Wilson loops. Our calculation is based on 500–700 configurations separated by 100–200 pseudo heat-bath sweeps. The smearing technique is used to reduce errors in the extracted potentials.

By construction, the canonical demon method preserves the average value of the Wilson loops included in the action. This means that in our case the average values of \( 1 \times 1 \) and \( 1 \times 2 \) Wilson loops \( \text{2} \), which are associated with \( \beta_{11} \) and \( \beta_{12} \), should be the same values as the blocked ones within statistical uncertainty. Fig. 2 and Fig. 3 show comparisons of \( 1 \times T \) Wilson loops between the original blocked lattices and the lattices obtained with the actions \( \text{A} \) and \( \text{B} \). As seen in both figures, good agreement is observed for each of \( 1 \times 1 \) and \( 1 \times 2 \) Wilson loops. This check ensures that the canonical demon method worked correctly. For other loops, however, the difference increases as \( T \) increases.

In order to compare the potentials we plot the potentials of the action \( \text{A}(\text{B}) \) and original blocked lattice on the same figure (Fig. 4, Fig. 5). It is clearly seen that the potentials of the action \( \text{A}(\text{B}) \) are very different from the ones of the original blocked lattice, and the slopes of the potentials at large distance of the action \( \text{A}(\text{B}) \) are

---

\( ^2 \) Some average values of Wilson loops of the MCRG actions are listed in Table 2.
larger, which implies that the lattice spacings are also larger than the original ones. We quantify this effect by comparing string tensions.

The string tensions are extracted by fitting the potentials to the form

$$V_L(r) = m + \sigma_L r - \frac{c}{r}$$  \hspace{1cm} (7)

where $m$, $\sigma_L$, and $c$ are fitting parameters, and $r = R/a$ is the distance measured in lattice units. The fits are performed on all the data including on and off axis potentials. The fit results are summarized in Table 3.
Fig. 4. Heavy quark potential for action A and the original lattice blocked at $\beta = 6.00$. The circles (diamonds) indicate on (off)-axis potentials. The curve in the figure is obtained by a fit to Eq. (7).

Fig. 5. Same as in Fig. 4 but for action B and the original lattice blocked at $\beta = 6.30$.

Table 3
Results of the fits. The lattice spacing $a$ is obtained by using $\sigma_L = a a^2$ and $\sqrt{\sigma} = 427$ MeV.

<table>
<thead>
<tr>
<th>Action</th>
<th>$m$</th>
<th>$\sigma_L$</th>
<th>$c$</th>
<th>$\chi^2$/d.o.f.</th>
<th>$a$(fm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.11(29)</td>
<td>1.29(12)</td>
<td>0.27(16)</td>
<td>1.37</td>
<td>0.52(5)</td>
</tr>
<tr>
<td>M</td>
<td>0.35(22)</td>
<td>0.885(85)</td>
<td>0.33(14)</td>
<td>0.42</td>
<td>0.43(4)</td>
</tr>
<tr>
<td>B</td>
<td>0.39(7)</td>
<td>0.587(27)</td>
<td>0.251(40)</td>
<td>1.17</td>
<td>0.35(2)</td>
</tr>
</tbody>
</table>
In order to evaluate the truncation effect, let us compare the string tensions with the ones of the original blocked lattices. The results are the following:

\[
\sigma_L = \begin{cases} 
1.29(12) & \text{MCRG Action A} \\
0.58(4) & \text{Original Blocked Lattice (at } \beta = 6.00) 
\end{cases}
\] (8)

\[
\sigma_L = \begin{cases} 
0.587(27) & \text{MCRG Action B} \\
0.287(26) & \text{Original Blocked Lattice (at } \beta = 6.30) 
\end{cases}
\] (9)

Here we comment on the naive expected string tension on a blocked lattice at \( \beta = 6.00 \) and 6.30. From the literature [14], we find \( \sigma a^2 \) at \( \beta = 6.00 \) is 0.051 and \( \sigma a^2 \) at \( \beta = 6.30 \) is 0.020. Since we block twice, the naive expected string tension on the blocked lattice will be \( 2 \sigma a^2 = 0.82(0.32) \) at \( \beta = 6.00 \) and \( \beta = 6.30 \) respectively. These values are compatible with that of the blocked lattice at \( \beta = 6.30 \) but not compatible at \( \beta = 6.00 \) (See Eqs. (8) and (9)). Probably this mismatch is due to the very small correlation length at \( \beta = 6.00 \) since the correlation length \( \xi \) is estimated to be \( \xi \sim 4 \) using \( 1/\xi^2 = \sigma a^2 \) and this correlation length is too small to preserve the same long range physics under the blocking transformation.

We examine the truncation effects which are seen in the string tension by taking the ratio of the string tensions of the MCRG actions and the original blocked lattices Eqs. (8), (9).

\[
\frac{(\sigma_L)^{\text{MCRG}}}{(\sigma_L)^{\text{Blocked}}} = \begin{cases} 
2.22 & \text{MCRG Action A} \\
2.04 & \text{MCRG Action B} 
\end{cases}
\]

Since \( \sigma_L = \sigma a^2 \), the ratio in lattice spacing will be

\[
\frac{(a)^{\text{MCRG}}}{(a)^{\text{Blocked}}} = \begin{cases} 
1.5 & \text{MCRG Action A} \\
1.4 & \text{MCRG Action B} 
\end{cases}
\] (10)

If the truncation effects were negligible, the ratio should be one. However this is far from being the case, which indicates that sizeable truncation effects are involved in the MCRG actions. It is important to notice here that the truncation effect increases the lattice spacing, i.e. the lattice spacing of the MCRG actions is 40–50% bigger than that of the original blocked lattice. This increase can be understood in the following way. The MCRG actions contain only the coupling constants associated with small Wilson loops \((1 \times 1 \text{ and } 1 \times 2 \text{ Wilson loops})\). On the other hand, the real action of the blocked configurations can have many coupling constants associated with large Wilson loops. Let us assume that all the long range coupling constants are positive. This assumption seems to be valid at least for coupling constants up to eight-links Wilson loop operators as found in [8], where all those coupling constants are positive. The positive coupling constants constrain Wilson loops to be more ordered. Since the large Wilson loops of the MCRG action are less ordered average values of the Wilson loops decay faster with distance than that of the constrained Wilson loops on the blocked configurations, as seen in Fig. 2, Fig. 3. Thus the potential of the MCRG action, \( V(R)^{\text{MCRG}} \), will be larger than that of the blocked configurations, \( V(R)^{\text{block}} \):

\[
V(R)^{\text{MCRG}} > V(R)^{\text{block}}.
\] (11)

This situation may become more pronounced at large \( R \) as seen in Fig. 4, Fig. 5. Thus, we obtain a larger string tension as extracted from the slope of the potential at large distance.

In order to examine the rotational symmetry of the potentials we plot all the data of the MCRG action A, B and M on the same figure (Fig. 6) by rescaling them using the obtained string tension results. The potentials are shifted so that each fitted curve for \( V(R)/\sqrt{\sigma} \) gives the value 2 at \( R/\sqrt{\sigma} = 2 \). We see good rotational symmetry except at large \( R/\sqrt{\sigma} \) due to the large error bars. The recovery of rotational symmetry will be more remarkable if

---

\footnote{We interpolate the value of the string tension for \( \beta = 6.30 \), since we do not find any Monte Carlo result at \( \beta = 6.30 \).}
we compare the potentials with that of the standard Wilson action \(^4\) where rotational symmetry is severely violated on a coarse lattice.

4. Discussion

The MCRG improvement approach has severe difficulties in determining the effective action. When we determine the action, we have to truncate it to a certain local form. On the other hand the truncation effects may ruin some benefits from the MCRG improvement. In order to reduce the truncation effects, one can tune the blocking transformation in such a way that contributions of long range coupling constants disappear quickly. This tuning was successfully done for the fixed point action \(^5\). At finite \(\beta\), however, a practical way to tune the blocking transformation is not known. The brute force approach, i.e. to directly search for an optimal blocking scheme by varying a couple of blocking parameters in Monte Carlo simulations, would certainly require a huge computational effort which we try to avoid.

Our MCRG improved actions have been obtained with a blocking scheme optimized by a Gaussian random SU(3) matrix. Strictly speaking, however, this optimization does not guarantee that the truncation effect will be minimized. Originally this optimization was introduced for the study of the \(\Delta \beta\), and the optimization was done so that the blocked trajectory converges to the renormalized trajectory quickly. This quick convergence is only desirable for the matching method of the \(\Delta \beta\) analysis. In this sense it is not sure that our MCRG actions are characterized by a local form. Actually our study found a big difference in the lattice spacing by the analysis of the string tension. Even so good news were found: one of the evidences for improvement, recovery of the rotational symmetry, is observed even in the presence of the truncation effects.

The one-loop level tadpole improvement scheme \(^4\) is known to work well, in which the action is designed to remove the \(\mathcal{O}(\alpha^2)\) and \(\mathcal{O}(\alpha_t \alpha^2)\) errors by adding two additional operators to the standard Wilson action. As far as rotational symmetry is concerned, the tree-level tadpole improvement scheme (Symanzik action + tadpole) also works well, in which one additional operator \((1 \times 2 \text{ Wilson loop})\) is added to the standard Wilson action. The only difference between the tree-level improvement (Symanzik action) with and without the tadpole

\(^4\) For example see Fig. 2 in [8] for comparison.
scheme is that the tadpole scheme increases the negative coupling contribution of the $1 \times 2$ Wilson loop operator, i.e., if we use the average plaquette value $\langle \text{plaq} \rangle = 0.4$, $\beta_{12}/\beta_{11}$ is approximately equal to $-0.08$, instead of $-0.05$ for the tree-level improvement scheme without tadpole improvement. For our MCRG actions, we find $\beta_{12}/\beta_{11} = -0.1$. The Iwasaki action, with which good rotational symmetry can be seen [15], also has similar behavior, i.e., $\beta_{12}/\beta_{11} = -0.09$. Therefore we suggest that rotational symmetry can be easily recovered with a short-distance correction ($1 \times 2$ Wilson loop) to the standard Wilson action and the value of $\beta_{12}/\beta_{11}$ should be more negative than the naive tree-level value, i.e., $\beta_{12}/\beta_{11} = -0.08$ to $-0.1$ in a region where the lattice spacing $a$ is as large as $a = 0.3$ to 0.5 fm.

The MCRG improved actions were determined with the canonical demon method. One could try another scheme such as the Schwinger-Dyson (SD) equation method [16]. Truncation effects with the SD equation method might be different, since in that case coupling constants are obtained by solving a set of equations whose coefficients are calculated from correlations among arbitrary Wilson loops, while the canonical demon method obtains coupling constants so that average values of Wilson loops included in the action only are conserved. And even though the canonical demon method conserves the average values of such Wilson loops, truncation effects already appear in the correlation among them [11]. Thus the SD equation might be more advantageous than the canonical demon method since the SD equation uses additional information on the action from the correlation among Wilson loops and determines the coupling constants by using many available equations. For the SU(3) gauge theory, determination of the renormalized trajectory was attempted with the SD equation method [17] and a rough picture of the RT was obtained. It would be interesting to see how truncation effects appear in a different way for both the canonical demon and the Schwinger-Dyson equation methods.

Acknowledgements

We thank the QCD-TARO Collaboration for their blocked configurations used in this study. T.T. would like to thank the Swiss Center for Scientific Computing, especially Professor Martin H. Gutknecht, for its hospitality.

References

First order phase transition in finite density QCD using the modulus of the Dirac determinant

R. Aloisio a,d , V. Azcoiti b , G. Di Carlo c , A. Galante a,c , A.F. Grillo d

a Dipartimento di Fisica dell’Università dell’Aquila, L’Aquila 67100, Italy
b Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain
c Istituto Nazionale di Fisica Nucleare, Laboratori Nazionali di Frascati, P.O.B. 13, Frascati 00044, Italy
d Istituto Nazionale di Fisica Nucleare, Laboratori Nazionali del Gran Sasso, Assergi (L’Aquila) 67010, Italy

Received 10 February 1998
Editor: L. Alvarez-Gaumé

Abstract

We report results of simulations of strong coupling, finite density QCD obtained within a MFA inspired approach where the fermion determinant in the integration measure is replaced by its absolute value. Contrary to the standard wisdom, we show that within this approach a clear signal for a first order phase transition appears with a critical chemical potential in extremely good agreement with the results obtained with the Glasgow algorithm. The modulus of the fermion determinant seems therefore to preserve some of the relevant physical properties of the system. We also analyze the dependence of our results on the quark mass, including both the chiral and large mass limit, and the theory in the quenched approximation.

Non-perturbative investigations of QCD at finite temperature and density have received much attention in the last years. The aim of these investigations is to find the matter conditions in the early Universe and to get a clear insight into experimental signatures in the heavy-ion collision experiments. Even if considerable progress has been achieved in the investigations of QCD at finite temperature and zero chemical potential using the lattice approach, the present situation of the field at finite density is not so satisfactory. As is well known, the complex nature of the determinant of the Dirac operator at finite chemical potential, which makes it impossible to use standard simulation algorithms based on positive-definite probability distribution functions, has much delayed investigations on the full theory with dynamical fermions. On the other hand the quenched approximation, which has been extensively and successfully used in simulations of QCD at zero chemical potential, seems to have some pathological behaviour when applied to QCD at finite density [1,2].

We discuss in this paper some features connected to simulations in finite density QCD. The main topic concerns the use of the absolute value of the fermionic determinant. We show that, against some theoretical prejudices based on random matrix models [3], the relevant physical features of finite density QCD seem to be preserved after taking the absolute value of the Dirac-Kogut-Susskind operator.

As well known, the partition function of finite density QCD

\[ Z = \int [dU] e^{-\beta S_{d}(U)} \det \Delta(U;m,\mu). \]
can be written as the product of the following two contributions

\[ Z = \langle e^{i\phi_2} \rangle \int \left[ dU \right] e^{-\beta S_G(U)} \left| \det \Delta(U;m,\mu) \right|. \tag{2} \]

where the first factor in (2)

\[ \langle e^{i\phi_2} \rangle = \frac{\int \left[ dU \right] e^{-\beta S_G(U)} e^{i\phi_2} \left| \det \Delta(U;m,\mu) \right|}{\int \left[ dU \right] e^{-\beta S_G(U)} \left| \det \Delta(U;m,\mu) \right|} \tag{3} \]

accounts for the mean value of the cosine of the phase of the fermion determinant computed with the probability distribution function of the pure gauge theory times the modulus of the fermion determinant. The second factor of (2) is just the partition function we will use along this work. The first factor in (2) gives a net contribution to the free energy density only in the case in which it falls off exponentially with the lattice volume. For random matrix models it has been shown that this is what happens [3] and this is the origin for the theoretical prejudices about the relevance of the phase in QCD. Early simulations of QCD with the absolute value of the fermion determinant in small lattices [4] seem to corroborate these theoretical prejudices.

For the abelian model in 0 + 1 dimensions it can be shown also that the first factor of (2) plays a fundamental role [5] and this is also the case for four dimensional QED as follows from the fact that the abelian model shows no dependence on the chemical potential \( \mu \) or in other words, from the absence of baryons in this model. Notwithstanding that, it can be shown that the phase of the determinant is completely irrelevant in 0 + 1 dimensional QCD [6] but unfortunately no analytical results on this subject are available for four dimensional QCD.

In order to check to what extent taking the absolute value of the fermion determinant in the integration measure is a good approximation for full QCD at finite baryon density, we report here results for the number density as a function of the chemical potential \( \mu \) at infinite gauge coupling, where more data are available in the literature [7,8]. The results for larger \( \beta \), in particular in the physically interesting scaling region, will be presented elsewhere.

Our numerical simulations have been performed using a MFA [9] inspired approach. The idea is to consider \( \det \Delta \) or its absolute value as an observable. In this case \( \det \Delta \) is not in the integration measure, and one avoids the problem of dealing with a complex quantity in the generation of configurations. This can be done in a (in principle) exact way by means of the MFA algorithm [9] where the mean value of the determinant at fixed pure gauge energy is used to reconstruct an effective fermionic action as a function of the pure gauge energy only. Up to now this method has been successfully used in several models (at zero density), where it allows free mobility in the \( \beta - m_q \) plane, including the chiral limit [9].

We used the GCPF (Grand Canonical Partition Function) formalism to write the fermionic determinant as a polynomial in the fugacity \( z = e^\mu [10]: \)

\[ \det \Delta(U;m_q,\mu) = \det(G + e^\mu T + e^{-\mu} T^\dagger) \]

\[ = z^{3V} \det( P(U;m_q) - z^{-1} ) \]

\[ = \sum_{n=-3L_z}^{3L_z} a_n z^n \]

where the propagator matrix is

\[ P(U;m_q) = \begin{pmatrix} -G & T \\ -T & 0 \end{pmatrix} \]

in which \( G \) contains the spatial links and the mass term, \( T \) contains the forward temporal links [10] and \( V \) is the lattice volume. Once fixed \( m_q \), a complete diagonalization of the \( P \) matrix allows to reconstruct \( \det \Delta \) for any \( \mu \). Due to the \( Z(L_z) \) symmetry of the eigenvalues of \( P \) it is possible to write \( P^{L_z} \) in a block matrix form and we only need to diagonalize a \((6L_z \times 6L_z)\) matrix; the chiral limit is straightforward since it only consists in diagonalizing \( P(U;m_q = 0) \).

Alternatively, one could directly diagonalize the fermionic matrix, expressing the determinant as a polynomial in the mass. In this way one is allowed to freely move in the fermionic mass, and this can be useful for the determination of chiral observables, but simulations have to be repeated for each value of the chemical potential. Also, this method tends to be more computer demanding, since it requires the diagonalization of larger matrices.
The partition function (1) is real, and positive definite; in particular, the average fermionic determinant at fixed energy is positive definite. This is no longer the case even in large but finite statistics: while the average determinant can be made real, its sign is not definite [4]. Therefore, and according to the previous discussion, we have chosen to compute $\mathcal{Z}$ by taking the absolute value of the fermionic determinant on a configuration by configuration basis. With the available statistics also other definitions we used to obtain a positive definite partition function like the modulus of the averaged coefficients or the modulus of the real part of the coefficients, instead of the average of the modulus gives results that can not be distinguished among them and with the previous definition.

We have performed simulations in $4^4$, $6^3 \times 4$ and $8^3 \times 4$ lattices, at $\beta = 0$ with various values of fermion masses, starting from $m_s = 0$. The analysis we report in this letter only concerns the behaviour of the baryonic number density $N(\mu)$ and that of its derivative with respect to $\mu$, which in the thermodynamical limit is proportional to the radial density of the zeros of the partition function. A more comprehensive analysis, reporting in particular the behaviour of chiral parameters is under way and will be published elsewhere.

The results for $\beta = 0$, $m_s = 0.1$, in $4^4$ and $6^3 \times 4$ lattices are reported in Figs. 1, 2. The values of the parameters have been chosen to allow a direct comparison with [7,8]. Several comments are in order.

1. In the smaller volume, neither $N(\mu)$ nor the radial density indicate a critical behaviour, apart from the (unphysical) behaviour near the onset threshold $\mu_0 = 0.31$. The small $\mu$ behaviour is unchanged in the larger volume, but a structure develops at $\mu_c = 0.69$ as well as at $\mu_c = 0.96$, which indicates a phase transition in the same position as that found by Karsch et al. [8] as well as a (first order) saturation transition, for which we lack of physical explanations. This contradicts the hypothesis [11] that taking the modulus of the determinant washes out the transition.

2. Our results are in very good agreement with others obtained with different methods. In particular, up to $\mu_c$ we are in striking agreement with Barbour et al. [7], reproducing both the “unphysical” behaviour at small $\mu$ and the transition at $\mu_c$. The value of $\mu_c$ agrees with that obtained in [7], but our discontinuity is steeper, corresponding to a higher peak in the radial distribution of zeros (see Fig. 2).

3. The value $\mu_0$, at which $N(\mu)$ departs from zero coincides with that found in [7]. Contrary to $\mu_c$, however, the peak at $\mu_0$ of the radial density of zeros does not drastically increase with the volume, as expected for a critical behaviour. We tend to exclude a first order phase transition at $\mu_0$.

4. The saturation at $N(\mu) = 1$ is not reached smoothly in the larger volumes, indicating a tran-

Fig. 1. Number density vs. chemical potential in a $4^4$ (dots) and $6^3 \times 4$ (diamonds) lattice at $m_s = 0.1$.

Fig. 2. Radial density of zeros in a $4^4$ (dots) and $6^3 \times 4$ (continuous line) lattices at $m_s = 0.1$. 
sition at $\mu_c = 0.96$. Also the radial density of zeros reported in [7] shows a peak consistent with the existence of this transition.

The behaviour sketched above pertains for values of fermion mass as low as 0.02. At lower masses the behaviour of the observables becomes smooth in $4^4$ and $6^3 \times 4$ lattices. In Fig. 3 we report preliminary results of a simulation at $m_q = 0, \beta = 0$ in a $8^3 \times 4$ for the radial density of the zeros superimposed with the same data for the smaller lattices. In this lattice a clear signal of transition develops at $\mu_c = 0.65$ as well as near saturation. This behaviour strongly suggests that the claim [11], that taking the modulus of the fermion determinant destroys the transition, is unjustified and what is observed is indeed a volume effect, as function of the fermionic mass, disappearing consistently in larger lattices. It is interesting to notice that (although more statistics is needed) the value of $\mu_c$ here obtained agrees with the chiral limit extrapolation of the value obtained in [8].

In Fig. 4 we report the dependence of $\mu_0$, $\mu_e$ and $\mu_s$ on the quark mass. One could expect that the partition function computed with the modulus of the determinant contains in the spectrum meson states with non zero baryonic charge, as in a theory with adjoint fermions in addition to usual ones. In this case one could expect a link between the features in $N(\mu)$ and the mass of the lightest of such states, degenerated with the ordinary pion [12]. In the figure we also plot the strong coupling/mean field value of half the mass of the pion [13]. Looking at the figure we can see that the onset $\mu_0$ follows rather closely the pion mass, at least for $m_q < 0.3$; for larger values of $m_q$ a clear displacement can be seen. Moreover the exact values of $\mu_0$ are better described, in the former region, by a simpler relation, i.e. the square root of the quark mass. If colourless diquark states, with non zero baryonic charge, lighter than the nucleon exist we would expect a saturation transition as soon as the chemical potential is of the order of magnitude of (half) the mass of the particle. Our data for $N(\mu)$ seem to discard this scenario; the (almost) linear behaviour between onset and the transition is not modified in the largest lattice.

We have also considered the quenched case where...
In conclusion, the situation is still far from clear. We have shown that using the modulus of the determinant we can reproduce the results obtained with other algorithms and this is the main motivation of this letter. However, the behaviour from the onset to the transition at \( \mu_c \) is difficult to understand in physical terms, as is the significance of the onset; we expect to get more informations from the analysis of chiral observables that we will present in a future publication. For large masses the technique based on the propagator matrix seems to suffer from numerical instabilities which may wash out the transition signal. On the other hand, the simulations performed diagonalizing directly the fermionic matrix appear to give results in agreement with physical intuition.

Acknowledgements

This work has been partly supported through a CICYT (Spain) – INFN (Italy) Collaboration. The authors thank Maria Paola Lombardo for useful discussions.

References

Right-handed currents in rare exclusive $B \to (K, K^*)\nu\bar{\nu}$ decays

D. Melikhov $^a$, N. Nikitin $^a$, S. Simula $^b$

$^a$ Nuclear Physics Institute, Moscow State University, Moscow 119899, Russia
$^b$ Istituto Nazionale di Fisica Nucleare, Sezione Sanita, Viale Regina Elena 299, I-00161 Roma, Italy

Received 13 March 1998

Abstract

The effects of possible right-handed weak hadronic currents in rare exclusive semileptonic decays $B \to (K, K^*)\nu\bar{\nu}$ are investigated using a lattice-constrained dispersion quark model for the calculation of the relevant mesonic form factors. The results obtained for the branching ratios and the missing energy spectra are presented and the sensitivity of various observables to long-distance physics is investigated. It is shown that the asymmetry of transversely polarized $K^*_T$ mesons as well as the $K/K^*$ production ratio are only slightly sensitive to long-distance contributions and mostly governed by the relative strength and phase of right-handed currents. In particular, within the Standard Model the production of right-handed $K^*_T$ mesons turns out to be largely suppressed with respect to left-handed ones, thanks to the smallness of the final to initial meson mass ratio. Therefore, the measurement of produced right-handed $K^*_T$ mesons in rare $B \to K^*\nu\bar{\nu}$ decays offers a very interesting tool to investigate right-handed weak hadronic currents.

PACS: 13.20.He; 12.39.Ki; 12.60.-i

Keywords: Semileptonic decays of bottom mesons; Relativistic quark models; Extensions of the Standard Model

Weak decays induced by Flavor Changing Neutral Currents (FCNC) are widely recognized as a powerful tool to make stringent test of the Standard Model (SM) as well as to probe possible New Physics (NP) [1]. Our understanding of FCNC in terms of the SM and its possible extensions is expected to be improved by the foreseen advent of new accelerators and $B$-factories, which will allow to investigate rare $B$-meson decays induced by the $b \to s$ (and $b \to d$) transitions. Besides many interesting processes, like, e.g., the $B_s - \bar{B}_s$ mixing and the rare $B$-meson decays induced by the $b \to s\gamma$ and $b \to s\ell^+\ell^-$ processes, the rare semileptonic decay $b \to s\nu\bar{\nu}$ plays a peculiar role. Indeed, within the SM the process $b \to s\nu\bar{\nu}$ is governed by the following effective weak Hamiltonian (cf. Refs. [2–4])

$$\mathcal{H}_{\text{eff}}(b \to s\nu\bar{\nu}) = \frac{G_F}{\sqrt{2}} \frac{\alpha_{\text{em}}}{2\pi\sin^2\theta_W} V_{tb} V_{ts}^* X(x_s) O_L(b \to s\nu\bar{\nu}) \equiv c_{L}^{(\text{SM})} O_L(b \to s\nu\bar{\nu})$$

where $O_L(b \to s\nu\bar{\nu}) \equiv (\bar{s}\gamma\gamma(1 - \gamma_5)b)(\bar{\nu}\gamma\gamma(1 - \gamma_5)\nu)$. The operator (1) is obtained from Z-penguin and W-box...
diagrams with a dominating top-quark intermediate state. In Eq. (1) $G_F$ is the Fermi constant, $\alpha_m$, the fine structure constant, $\theta_W$ the Weinberg angle, $V_{tb}/V_{ts}$ the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements and $x_t = (m_t/m_W)^2$; finally, the function $X(x_t)$ is obtained after integrating out the heavy particles and includes $O(\alpha)$ corrections (see Refs. [2,3] for its explicit expression).

The appealing feature of Eq. (1) relies in the presence of a single operator governing the transition $b \rightarrow s \nu \bar{\nu}$. In this way the main theoretical uncertainties are concentrated in the value of only one Wilson coefficient, $c_L^{\text{SM}}$. As it is well known, in case of other processes, like those driven by the $b \rightarrow s \gamma$ and $b \rightarrow s \ell^+ \ell^-$ transitions, several operators should be included in the effective weak SM Hamiltonian, so that the corresponding set of Wilson coefficients (with their theoretical uncertainties) act coherently in determining the values of many observables, like the branching ratios, the differential decay rates, lepton asymmetries, etc. Moreover, the SM operator (1) does not contain long-distance contributions generated by four-quark operators, which are usually present in the low-energy weak Hamiltonian and affect both the $b \rightarrow s \ell^+ \ell^-$ and (to a much less extent) the $b \rightarrow s \gamma$ processes.

Under the only assumption of purely left-handed neutrinos (possible neutrino mass effects are expected to be negligible [5]) NP effects in the $b \rightarrow s \nu \bar{\nu}$ transitions can be a modification of the SM value of the coefficient $c_L$ and/or the introduction of a new right-handed (RH) operator, viz.

$$\mathcal{H}_{\text{eff}}(b \rightarrow s \nu \bar{\nu}) = c_L O_L(b \rightarrow s \nu \bar{\nu}) + c_R O_R(b \rightarrow s \nu \bar{\nu})$$

(2)

where $O_R(b \rightarrow s \nu \bar{\nu}) \equiv (\bar{\nu}_R(1 + \gamma_5)b)(\bar{\nu}_R(1 - \gamma_5)b)$ and the values of the coefficients $c_L$ and $c_R$ depend on the specific NP model. In what follows, we will make use of two parameters, $\epsilon$ and $\eta$, defined as

$$\epsilon^2 = \frac{|c_L|^2 + |c_R|^2}{|c_L^{\text{SM}}|^2}, \quad \eta = -\frac{\text{Re}(c_L^* c_R)}{|c_L|^2 + |c_R|^2}$$

(3)

which are clearly connected to the relative strength and phase of RH currents with respect to left-handed (LH) ones.

From the theoretical point of view the inclusive $B \rightarrow X_s \nu \bar{\nu}$ decay is a particularly clean process for investigating possible NP effects, because the non-perturbative $1/m_b^2$ corrections to the free-quark result are known to be small [6]. This is valid not only for the branching ratio, but also for the differential decay rate, except for the regions close to the kinematical end-point, where the spectrum has to be smeared out to get reliable results. The free-quark prediction for the missing-energy spectrum of the $B \rightarrow X_s \nu \bar{\nu}$ decay can be read off from, e.g., Ref. [5]. While the differential decay rate is directly proportional to $\epsilon^2$, its dependence upon $\eta$ is due to interference effects between the LH and RH currents. However, since the final to initial quark mass ratio, $m_b/m_{u,s}$, is quite small, the shape of the missing energy spectrum is only slightly affected by the value of $\eta$ (cf. Ref. [5]). Therefore, a full determination of the coefficients $c_L$ and $c_R$ requires at least to consider also decay processes other than the inclusive $B \rightarrow X_s \nu \bar{\nu}$ one. It is the aim of this letter to show that the effects of possible RH currents can be investigated in rare exclusive semileptonic decays $B \rightarrow (K, K^*) \nu \bar{\nu}$ and to this end a lattice-constrained dispersion quark model, recently developed in Ref. [7], is used to evaluate the relevant mesonic form factors. It will be shown that the asymmetry of transversely polarized $K^*_2$ mesons as well as the $K^*/K_{2}^*$ production ratio are only slightly affected by the model dependence of the form factors and remarkably sensitive both to $\epsilon$ and $\eta$, i.e. to the relative strength and phase of RH currents. In particular, within the SM the production of RH $K^*_2$ mesons turns out to be largely suppressed with respect to LH ones, thanks to the

\[1\] It turns out that the main uncertainty on $c_L^{\text{SM}}$ is the uncertainty on the top-quark mass and on the product $|V_{tb}V_{ts}^*|$ of CKM matrix elements [4], being the radiative QCD corrections substantially small [3].
smallness of the final to initial meson mass ratio and, therefore, the measurement of produced RH $K^+_\tau$ mesons in rare $B \to K^+ \nu \bar{\nu}$ decays offers a very interesting tool to investigate RH weak hadronic currents.

To begin with, let us denote by $P_\mu$ and $P_{K(K^+)}$ the four-momentum of the initial and final mesons and define $q = P_\nu - P_{K(K^+)}$ as the four-momentum of the $\nu \bar{\nu}$ pair and $x \equiv E_{\text{miss}}/M_B$ the missing energy fraction, which is related to the squared four-momentum transfer $q^2$ by: $q^2 = M_B^2 \left(2x - 1 + r_{K(K^+)}^2\right)$, where $r_{K(K^+)} \equiv M_{K(K^+)}/M_B$ with $M_B$ and $M_{K(K^+)}$ being the initial and final meson masses. The missing energy spectrum for the decay $B \to K^+ \nu \bar{\nu}$ can be written as (cf. Refs. [7,8])

$$\frac{dBr}{dx} (B \to K^+ \nu \bar{\nu}) = 3Br_0 \left[ \frac{c_L + c_R}{c_L^{\text{SM}}} \right]^2 \left(1 - x\right)^{2} - r_{K(K^+)}^2 \frac{3}{2} \left[F_i(q^2)\right]^2$$

(4)

where the factor 3 arises from the sum over the three neutrino generations and

$$Br_0 \equiv \rho_{c_L^{\text{SM}}} \frac{M_B^2 \tau_B}{4\pi^2} \frac{G_F^2 P^2}{M^2} \left(4\pi\right)^2 \frac{4\alpha^2}{\sin^4(\theta_C)} X^2(x_i) |V_{ub}V_{us}^*|^2 \tau_B$$

(5)

with $\tau_B$ being the $B$-meson lifetime. In Eq. (4) the mesonic form factor $F_i(q^2)$ is obtained from the covariant decomposition of the hadronic transition $B \to K$ driven by the vector current $\gamma_\mu b$, viz.

$$\langle K | \gamma_\mu b | B \rangle = \left(P_\mu + P_{K(K^+)}\right) \cdot F_1(q^2) + q_\mu \left[F_0(q^2) - F_1(q^2)\right] \frac{M_B^2 - M_{K(K^+)}^2}{q^2}$$

(6)

with $F_0(q^2 = 0) = F_1(q^2 = 0)$. As for the $B \to K^+ \nu \bar{\nu}$ decay, the missing energy spectrum corresponding to a definite polarization $\gamma (= 0, 1)$ of the final $K^+$ meson is given by [8]

$$\frac{dBr}{dx} (B \to K^+ \nu \bar{\nu}) = \frac{1}{3} Br_0 \left[ \frac{c_L + c_R}{c_L^{\text{SM}}} \right]^2 \frac{1}{r_K^2 \cdot (1 + r_{K(K^+)}^2)} \left(1 - x\right)^{2} - r_{K(K^+)}^2 \cdot$$

$$\times \left(1 + r_{K(K^+)}\right) \left(1 - x - r_{K(K^+)}^2\right) A_1(q^2) - 2 \left(1 - x\right)^{2} - r_{K(K^+)}^2 \cdot A_2(q^2)$$

(7)

$$\frac{dBr}{dx} (B \to K^+ \nu \bar{\nu}) = \frac{2}{3} Br_0 \left(1 - x\right)^{2} - r_{K(K^+)}^2 \cdot \frac{2x - 1 + r_{K(K^+)}^2}{1 + r_{K(K^+)}^2}$$

$$\times \left[ \frac{2}{c_L + c_R} \right] \left(1 - x\right)^{2} - r_{K(K^+)}^2 \cdot V(q^2) \left[ \frac{c_L - c_R}{c_L^{\text{SM}}} \right] \left(1 + r_{K(K^+)}^2\right) A_1(q^2)$$

(8)

where the mesonic form factors $V(q^2), A_1(q^2)$ and $A_2(q^2)$ appear in the covariant decomposition of the hadronic matrix elements of the $B \to K^+$ transition generated by the $V - A$ current $\gamma_\mu (1 - \gamma_5) b$, viz.

$$\langle K^+ | \gamma_\mu (1 - \gamma_5) b | B \rangle = e_{\mu \nu \alpha \beta} \varepsilon^\nu \varepsilon^\alpha (h) \left(P_\mu - P_{K^+}\right) \frac{2V(q^2)}{M_B + M_{K^+}}$$

$$- i \left[ e^\nu (h) \left(M_B + M_{K^+}\right) A_1(q^2) - \left[e^\alpha (h) \cdot q\right] \left(P_\mu + P_{K^+}\right) \cdot A_2(q^2) \right] \frac{2M_{K^+}}{q^2} \left[A_3(q^2) - A_0(q^2)\right]$$

(9)

where $\varepsilon$ is the polarization four-vector of the $K^+$-meson and $A_0(q^2) \equiv \left[(M_B + M_{K^+}) A_1(q^2) - (M_B - M_{K^+}) A_2(q^2)\right]/2M_{K^+}$. 
In the whole accessible kinematical decay region we have calculated the relevant mesonic form factors $F_q(q^2)$, $V(q^2)$, $A_q(q^2)$ and $A_s(q^2)$, appearing in Eqs. (4) and (7), (8), adopting a dispersion formulation of the relativistic quark model [9]. For the explicit evaluation of the form factors one needs to specify the quark model parameters such as the constituent quark masses and the meson wave functions. In Ref. [10] we performed calculations of the mesonic form factors adopting different model wave functions, in particular: the simple Gaussian ansatz of the ISGW2 model [11] and the variational solution [12] of the effective $q\bar{q}$ semi-relativistic Hamiltonian of Godfrey and Isgur (GI) [13]. These two models differ both in the shape of the meson wave functions, particularly at high internal momenta, and in the values of the quark masses (see Ref. [10] for details). The results of our calculations showed that the mesonic form factors for heavy-to-light transitions are sensitive both to the high-momentum tail of the meson wave function and to the values adopted for the quark masses (see also Ref. [14]). In order to obtain more reliable predictions for the form factors, we have required [7] the quark model parameters to be adjusted in such a way that the calculated form factors at high $q^2$ are compatible with recent available lattice QCD results [15,16]. We found that the best agreement with the high-$q^2$ lattice data can be obtained adopting the quark masses and wave functions of the GI model with a switched-off one-gluon exchange, which will be denoted hereafter to as the GI–OGE quark model.

Recently, in the whole range of accessible values of $q^2$ a lattice-constrained parametrization for the $B \to K^*$ form factors has been developed [17], based on the Stech’s parametrization of the form factors obtained within the constituent quark picture [18], on the Heavy Quark Symmetry (HQS) scaling relations near $q^2 = M_B^2 - M_{K^*}^2$ and on a single-pole behavior of $A_q(q^2)$ suggested by the heavy-quark mass dependence at $q^2 = 0$ expected from the QCD sum rules (see Ref. [19] for details). The parameters of the single-pole fit to the form factor $A_q(q^2)$ were found from the least-$\chi^2$ fit to the lattice QCD simulations [15,16] in a limited region at high values of $q^2$. Such a parametrization, though still phenomenological, is also consistent with the dispersive bounds of Ref. [19] and therefore it obeys all known theoretical constraints. The comparison of the mesonic form factors relevant in rare $B \to (K, K^*)\pi\bar{\pi}$ decays, obtained in our GI–OGE relativistic quark model, with the parametrization of Ref. [17] is shown in Fig. 1. It can clearly be seen that the two sets of form factors agree each other within $\approx 10\%$, except near the zero-recoil point. We want to stress that both sets of form factors satisfy all known rigorous theoretical constraints in the whole kinematical accessible region. Therefore, we expect that the difference between the two sets of form factors provide a typical present-day theoretical
uncertainty of our knowledge of long-distance effects in the mesonic channels. Consequently, in order to investigate the sensitivity of the missing-energy spectra and branching ratios of rare $B \to (K, K^*) \ell \ell$ decays to the specific $q^2$-behavior of the relevant form factors, we have run calculations of Eqs. (4) and (7), (8) adopting the two sets of mesonic form factors shown in Fig. 1. We have also adopted the following values: $\tau_B = (1.57 \pm 0.04)$ ps [20], $M_B = 5.279$ GeV [21], $\sin(\theta_W) = 0.2315$ [21] and $|V_{tb}V_{ts}^*| = 0.038 \pm 0.005 (0.041 \pm 0.005)$ in case of our GI–OGE form factors (parametrization of Ref. [17]). Finally, at $m_t = 176$ GeV one gets $X(x_t) = 2.02$, yielding $Br_{B} = (5.3 \pm 1.4) \times 10^{-4}$ (see Eq. (5)), which implies a present-day theoretical uncertainty of ~25% in rare $b \to s \ell \ell$ decay rates.

The present experimental upper bound on the inclusive branching ratio $Br(B \to X_s \ell \ell)$, determined by the ALEPH collaboration ($Br(B \to X_s \ell \ell) < 7.7 \times 10^{-4}$ [22]), turns out to be about an order of magnitude larger than the typical SM prediction ($Br_{SM} \to X_s \ell \ell = 4 \times 10^{-5}$ [3,8]). Thus, till now the inclusive $B \to X_s \ell \ell$ decay constrains weakly the range of values of the relative strength factor $e^2$, namely: $e^2 < 15 \cdot Br(B \to X_s \ell \ell)/(7.7 \times 10^{-4})$ and, in general, the admixture of possible RH currents in $b \to s$ transitions is not yet constrained too much, leaving the possibility of a sizable strength with unknown relative phase (see, e.g., [23]). Since the present SM uncertainty on rare $b \to s \ell \ell$ decay rates is about 25%, we have simply considered a relative strength factor $e^2$ equal to 1.25 (i.e., a 25% enhancement with respect to the SM value $\epsilon_{SM}^2 = 1$) and varied the parameter $\eta$ in its allowable range. Note that the value $e^2 = 1$ can be realized not only within the SM framework, but possibly also in NP scenarios; however, for sake of simplicity, in what follows we will assume $c_L = c_{L}^{SM}$. Our results obtained for the branching ratios $Br(B \to K \ell \ell)$, $Br(B \to K^* \ell \ell)$ and $Br(B \to K_{\ell}^{*+} \ell \ell)$ (where $K_{\ell}^{*+}$ and $K^*_{\ell}$ stand for longitudinally and transversely polarized $K^*$-mesons, respectively), are reported in Fig. 2(a) and clearly show that the production of longitudinally polarized $K_{\ell}^{*+}$-mesons is remarkably sensitive to the detailed $q^2$-behavior of the mesonic form factors, while our predictions for both $K_{\ell}^{*+}$ and $K^*$-meson production are only slightly model-dependent. Note that, contrary to what happens in case of the inclusive $B \to X_s \ell \ell$ process, the exclusive branching ratios $Br(B \to (K, K^*) \ell \ell)$ are much more sensitive to the value of the parameter $\eta$, i.e. to the relative phase of RH currents with respect to LH ones, even when a quite small

---

2 The values adopted for $|V_{tb}V_{ts}^*|$ for the two sets of form factors have been determined in Ref. [7] by the request of reproducing the CLEO result [24] on the rare exclusive $B \to K^* \ell \ell$ decay within the SM basis.
Fig. 3. Differential branching ratio of the decay process $B \to K^*\nu\bar{\nu}$ (see Eqs. (7), (8)) versus the missing-energy fraction $x$, calculated within the SM framework. The dot-long dashed, dashed and solid lines correspond to final $K^*$-mesons with polarization $h = 0, +1, -1$, respectively. The thin and thick lines have the same meaning as in Fig. 1.

enhancement factor $e^2 - 1$ is considered. Our predictions for the ratio $R_{K/\bar{K}}$ of produced $K$- to $\bar{K}$-mesons as well as the transverse asymmetry $A_T$, defined as

$$R_{K/\bar{K}} \equiv \frac{Br(B \to K\nu\bar{\nu})}{Br(B \to K^{*-} \nu\bar{\nu}) + Br(B \to K^{-} \nu\bar{\nu})},$$

$$A_T \equiv \frac{Br(B \to K^{*-} \nu\bar{\nu}) - Br(B \to K^{-} \nu\bar{\nu})}{Br(B \to K^{*-} \nu\bar{\nu}) + Br(B \to K^{-} \nu\bar{\nu})}$$

are shown in Fig. 2(b). It can be seen that the transverse asymmetry $A_T$ is only marginally affected by the model-dependence of the mesonic form factors, while the ratio $R_{K/\bar{K}}$ is more sensitive to their specific $q^2$-behavior, particularly at negative values of $\eta$. It should be pointed out however that both $A_T$ and (to a much larger extent) $R_{K/\bar{K}}$ are remarkably affected by the value of $\eta$. The advantage of such a large sensitivity should be taken into account when comparing inclusive versus exclusive decay modes, the latter being affected by the general problem of the model dependence of the form factors.

It can be easily checked (starting from Eqs. (4) and (7), (8)) that the ratio $R_{K/\bar{K}}$ is independent of $\epsilon$, while the asymmetry $A_T$ depends both on $\eta$ and $\epsilon$ \(^3\). Our predictions for the SM values of $R_{K/\bar{K}}$ and $A_T$ are 0.76 ± 0.04 (see Fig. 2(b) at $\eta = 0$) and 0.93 ± 0.02, respectively, where the quoted uncertainties correspond to the variation obtained using the two sets of mesonic form factors adopted in this work. The SM value of $A_T$ indicates a dominance of produced LH ($h = -1$) $K^*$-mesons with respect to RH ($h = +1$) ones. This fact is clearly illustrated in Fig. 3, where our predictions for the missing-energy spectra of longitudinally and transversely polarized $K^*$-mesons, obtained within the SM framework, are reported. The main outcome can be summarised as follows: i) the energy spectrum of longitudinally polarized $K^*$-mesons is largely affected by the model-dependence of the mesonic form factors, while the opposite feature is exhibited by the production of transversely polarized $K^*$-mesons, and ii) RH $K^*$-mesons are less abundant than LH ones within the SM in a wide range of values of $x$, except near the zero-recoil point (corresponding to $x = x_{\text{max}} = 1 - r_{K^*}$). We have reached the same conclusions also after having carried out the calculations of the transverse asymmetry $A_T$ using other sets of form factors, like those obtained within the two QCD sum rule versions of Refs. [8] and [25],

\(^3\) Note also that both $R_{K/\bar{K}}$ and $A_T$ are clearly independent of $|V_{td}V_{ts}^*|$ and the top-quark mass.
or the form factors fulfilling the HQS relations at leading-order in the inverse heavy-quark mass (see Refs. [26, 10]). As for the latter case, it is well known that in the heavy-quark limit (HQL) all the relevant form factors are related to a single universal function, the Isgur-Wise form factor. Moreover, in the HQL the (differential) transverse asymmetry $A_T(x)$ is independent of the Isgur-Wise function and within the SM it simply reduces to a kinematical function, viz:

$$A_T(x) = \frac{d\text{Br}(B \to K_{h=1}^{*+} \nu\bar{\nu})/dx - d\text{Br}(B \to K_{h=+1}^{*+} \nu\bar{\nu})/dx}{d\text{Br}(B \to K_{h=0}^{*+} \nu\bar{\nu})/dx + d\text{Br}(B \to K_{h=+1}^{*+} \nu\bar{\nu})/dx} \left( \frac{\sqrt{\omega^2 - 1}}{\omega} \right)$$

(11)

where $\omega$ is dot product of the initial and final meson four-velocities. Therefore, starting from the zero-recoil point ($\omega = 1$) where $A_T(x) = 0$, the transverse asymmetry in the SM rapidly increases up to its maximum value $(1 - r_K^2)/(1 + r_K^2)$, reached at the maximum-recoil point $\omega_{\max} = (1 + r_K^2)/2 r_K^2$. (corresponding to $x = x_{\max} = (1 - r_K^2)/2$). Since $r_K^2 = 0.03$ one has $A_T(x) \sim 0.9$ in a wide range of values of $x$. Though the final $K^*$-mesons are far from being considered as heavy daughters, approximate HQS relations among the form factors of the $B \to K^*$ transition have been shown to hold within $\pm 20\%$ accuracy [7, 10], so that the dominance of LH produced $K_T^*$ mesons within the SM holds not only in the HQL, but also in case of finite quark masses. To sum up, the measurement of produced RH $K_T^*$-mesons in rare $B \to K^* \nu\bar{\nu}$ decays could offer a clear signature of possible RH weak hadronic currents. Finally, we have collected in Fig. 4 our predictions for the shape of the missing-energy spectra of the $B \to K_{h=0}^{*+} \nu\bar{\nu}$ decay, obtained for various values of $\epsilon$ at fixed value of $\eta$ (see Fig. 4(a)) as well as for various values of $\eta$ at fixed $\epsilon$ (see Fig. 4(b)). Note in particular that the production of LH $K_T^*$-mesons is almost independent of the value of $\epsilon$, while the differential rate for RH ones is approximately proportional to $(\epsilon^2 - 1)$.

In conclusion, the missing-energy spectra and branching ratios of rare exclusive semileptonic $B \to (K, K^*) \nu\bar{\nu}$ decays have been investigated adopting a lattice-constrained dispersion quark model for the calculation of the relevant mesonic form factors. The effects of possible right-handed weak hadronic current have been considered and the sensitivity of the branching ratios and the missing energy spectra to long-distance physics has been investigated. It has been shown that the asymmetry of transversely polarized $K_T^*$ mesons as well as the $K/K_T^*$ production ratio are only slightly sensitive to long-distance contributions and mostly governed by the relative strength and phase of right-handed currents. In particular, within the Standard Model the production of right-handed $K_T^*$ mesons turns out to be largely suppressed with respect to left-handed ones, thanks to the

![Fig. 4. Missing-energy spectrum of transversely polarized $K_T^*$-mesons produced in rare $B \to K^* \nu\bar{\nu}$ decays, calculated using our lattice-constrained dispersion quark model. In (a) the parameter $\eta$ (Eq. (3)) is fixed at the value $\eta = 0$, while the solid, dashed, dotted and dot-long dashed lines correspond to $\epsilon = 1.0, 1.05, 1.25$ and 1.55, respectively. The thin and thick lines are the results obtained in case of RH and LH final $K_T^*$-mesons. In (b) the same as in (a), but at $\epsilon = 1.25$ for various values of $\eta$; the dashed, solid and dot-long dashed lines correspond to $\eta = -0.25, 0.0$ and 0.25, respectively.](image-url)
smallness of the final to initial meson mass ratio. Therefore, the measurement of produced right-handed $K^+_T$ mesons in rare $B \to K^+\nu\bar{\nu}$ decays offers a very interesting tool to investigate right-handed weak hadronic currents and, despite the general problem of the model dependence of the hadronic form factors, the exclusive decay modes $B \to (K, K^+)\nu\bar{\nu}$ turn out to be more sensitive to the effects of right-handed currents with respect to the inclusive $B \to X_s\nu\bar{\nu}$ process.

References

Isospin breaking and the extraction of $m_s$ from the $\tau$-decay-like vector current sum rule

Kim Maltman 1

Department of Mathematics and Statistics, York University, 4700 Keele St., North York, Ont. M3J 1P3, Canada
and Special Research Center for the Subatomic Structure of Matter, University of Adelaide, 5005 Australia

Received 22 October 1997; revised 27 February 1998

Abstract

Narison’s $\tau$-decay-like sum rule for determining the strange quark mass is re-investigated, taking into account isospin-breaking corrections in the extraction of the input spectral functions from $e^+ e^- \to$ hadrons data. The corrections, estimated using experimental data on vector meson electromagnetic decay constants and a QCD sum rule analysis of the $38$ vector current correlator, are shown to be especially large for the isoscalar case. The reason such large corrections are natural is also explained. Due to the high degree of cancellation in the original sum rule, the effect of these corrections on the determination of $m_s$ is significant. A new central value $m_s = 113 \pm 138$ MeV is found, in the $\overline{\text{MS}}$ scheme at 1 GeV, with significant (asymmetric) errors associated with errors in the input experimental data. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

The usually quoted values for the light $(u,d,s)$ current quark masses are obtained by a sum rule analysis of the correlator of either (for $m_s + m_d$) the product of two divergences of the isovector axial vector current $[1,2]$ or (for $m_s$) the product of two divergences of the strangeness-changing vector current $[3,4]$. One of the potential problems with this approach is that, in both cases, continuum contributions to the hadronic side of the sum rule are large, but the continuum part of the hadronic spectral function is not known experimentally. The extracted quark masses thus depend crucially for their reliability on that of the theoretical ansatze for the unmeasured continuum spectral functions, which are constructed by analogy with an extreme form of the vector meson dominance (VMD) treatment of the vector isovector channel wherein the continuum spectral function is modelled as a sum of Breit-Wigner terms whose overall normalization is adjusted to produce the desired value at continuum threshold. The threshold value is estimated, for the pseudoscalar channel, using tree-level chiral perturbation theory (ChPT) $[1]$, and for the $S = -1$ scalar channel, by extrapolating $K_S$ data using the Omnes representation, together with experimental data on the $K\pi$ scattering phases $[3]$. Since the relevant thresholds are many resonance widths away from the poles, the assumed $q^2$-dependence of the resonance widths is also a crucial input. Taking the $m_s$ analysis to be specific, the “standard” form of the $q^2$-depen-
dent $s$-wave width is employed. This form results from assuming the scalar couplings of the resonances in question to $K\pi$ are $q^2$-independent, a somewhat dangerous assumption in the scalar channel.

The above assumptions have generally been considered plausible because an analogous version is known to allow a successful description of $e^+e^- \to \pi^+\pi^-$ cross-sections. The analogy, however, potentially dangerous [5–7]. In Ref. [7], for example, the $K\pi$ portion of the scalar, $S = -1$ spectral function is obtained for all $s$ using the Omnès representation of the timelike scalar $K\pi$ form factor, with $K_{\pi,3}$ and $K\pi$ phase data as input. The resulting spectral function rises much faster just above threshold, and reaches a much lower (by a factor of $\sim 3$) $K_0^-$ (1430) peak height, than does the model version, obtained using an assumed resonant spectral shape normalized at threshold, employed in Refs. [3, 4].

Above $s \approx 2\text{ GeV}^2$, where $K\pi$ states no longer dominate the spectral function, a full determination of the relevant spectral function via the method of Ref. [7] is no longer possible; below $s \approx 2\text{ GeV}^2$, however, this result clearly demonstrates the existence of potentially large uncertainties in the "resonance-saturation/threshold-normalization" ansätze for the unmeasured continuum spectral functions. This suggests it is preferable to employ a sum rule involving experimentally-determinable spectral data. The sum rule for $m_s$ proposed by Narison [8] is of this type.

Narison’s idea is to consider the difference of isovector and hypercharge vector current correlators, $\Pi^{33} - \Pi^{88}$, where 3 and 8 are $SU(3)_F$ labels, and $\Pi^{aa}$ is defined by

$$i \int d^4x e^{iq\cdot x} \langle 0 | \mathcal{T} \left( J^{aa}_0(x) J^{aa}_0(0) \right) | 0 \rangle \equiv \left( q_\mu q_\nu - q^2 g_{\mu\nu} \right) \Pi^{aa}(q^2).$$

In the operator product expansion (OPE) for this difference, each term necessarily involves the flavor-breaking parameter $m_A - \bar{m}$ (where $\bar{m} = (m_u + m_d)/2$). Narison proposes integrating the corresponding spectral function, as weighted for $\tau$ decay kinematics, from threshold $q^2 = 4m^2_A$ up to a mock $\tau$ mass, $q^2 = m^2_\tau$. As for inclusive hadronic $\tau$ decays [9–14], analyticity properties allow one to re-write this integral as a correspondingly weighted integral of the correlator difference over a circular contour of radius $m^2_\tau$ in the complex $q^2$ plane. For large enough $m_\tau$, this alternate representation simultaneously suppresses both contributions from the region of the contour near the positive real axis where perturbative QCD (pQCD) becomes unreliable, and non-perturbative contributions in the OPE relative to perturbative ones [9–14]. In addition, if one ignores isospin breaking, the relevant spectral functions are experimentally determinable in $e^+e^- \to \tau^n\bar{\tau}$ hadrons. Narison’s original treatment [8], which neglected isospin breaking, produced $m_s$ values compatible with those of the conventional analyses [3, 4] (197 ± 29 MeV in the MS scheme, at a scale of 1 GeV$^2$, cf. 205 ± 19 MeV [4]). Since, however, there is a high degree of cancellation between the 33 and 88 contributions to the sum rule, it is important to consider the possibility of isospin-breaking corrections to the hadronic spectral functions. We investigate this question in the present paper.

2. Narison’s $\tau$-decay-like sum rule for $m_s$

Narison’s method is based on an analogy with analyses of the inclusive $\tau$ decay ratio

$$R^{l=1}_{\tau} = \frac{\Gamma_{\tau \to \ell \nu}^{\text{hadrons}}(\gamma)}{\Gamma(\tau \to e\nu_e\bar{\nu}_e(\gamma))},$$

(for concreteness, we consider $\tau$ decays mediated by the charged weak isovector vector current). $R^{l=1}_{\tau}$ is given by an integral over the $J = 0, 1$ scalar spectral functions of the vector isovector correlator, weighted by the appropriate kinematic factors [9–14], which integral can be converted into one involving the scalar correlators themselves, with the same kinematic weights, over the counterclockwise oriented circular contour of radius $m^2_\tau$ in the complex $s = q^2$ plane. Narison considers the difference $R^{33}_{\tau} - R^{88}_{\tau}$, where $R^{aa}_{\tau}$ ($a = 3, 8$) results from replacing $m_A$ by a variable mass, $m_{\tau}$, and the isovector current in $R^{l=1}_{\tau}$ by $J^{aa}_0$. This difference can be expressed in either the "hadronic" or "contour integral" representations

$$[R^{aa}_{\tau}]_{\text{had}} = 12\pi^2 |V_{ud}|^2 S_{\text{EW}} \int_0^{m^2_\tau} ds \mu^2 \left( 1 - \frac{s}{m^2_\tau} \right)^2$$

$$\times \left( 1 + \frac{2s}{m^2_\tau} \right) \rho^{aa}(s),$$

\[\]
$$\left[ R_{\mu}^{\alpha\beta} \right]_{\text{contour}} = 6\pi i |V_{ud}|^2 S_{\text{EW}} \int_{s=0}^{\infty} \frac{ds}{m_i^2} \left( 1 - \frac{s}{m_i^2} \right)^2$$

$$\times \left[ 1 + \frac{2s}{m_i^2} \right] \Pi^{\alpha\beta}(s),$$

(4)

where $V_{ud}$ is the $u\bar{d}$ CKM matrix element, $S_{\text{EW}} = 1.0194$ is the sum of the leading-log electroweak $\tau$ decay corrections [15] via the $s$- and isoscalar spectral functions and, for sufficiently large $m_i$, $\left[ R_{\mu}^{\alpha\beta} \right]_{\text{contour}}$ can be evaluated using the OPE. The sum rule for $m_i$ results from equating Eqs. (3) and (4) [8].

For the OPE side one has, with $D$ labelling operator dimension [8],

$$\left[ R_{\mu}^{33} - R_{\mu}^{s3} \right]_{\text{OPE}} = |V_{ud}|^2 S_{\text{EW}} \sum_{D=2,4,\ldots} \Pi^{\alpha\beta}(s),$$

(5)

The $D = 2$ terms result from the mass-dependent perturbative contributions to the correlators which, for a flavor-diagonal vector current of flavor $\nu$, are [16]:

$$\Delta_{\nu}^{(\text{mass})}(s) = -\frac{3}{2\pi} \frac{m_i(s)}{Q^2} \left[ 1 + \frac{s}{m_i^2} a(Q^2) + \frac{12}{13} \left( \frac{m_i^2}{Q^2} - 1 \right) \frac{a(Q^2)}{\alpha}(s) \right]$$

$$+ \frac{1}{\alpha}(s) \left( 32 - 24\frac{m_i^2}{Q^2} \right),$$

(6)

with $a(Q^2) = \alpha(Q^2)/\pi$ and $\delta_i(Q^2)$ the running coupling and running mass of quark $j$, both at scale $\mu_i^2 = Q^2 - s$, in the MS scheme. Expanding $a(Q^2)$ in terms of $m_i^2$ and $\delta_i(Q^2)$ in terms of $a$ and $\delta_i$ in terms of $m_i$, and performing the resulting elementary logarithmic integrals, one obtains

$$\delta_{\nu}^{(2)} - \delta_{s}^{(2)} = \frac{12 \left( m_i^2 - m_j^2 \right)}{m_i^2} \left( 1 + f_1 a + f_2 a^2 + f_3 a^3 + \cdots \right),$$

(7)

where $f_1 = 13/3$, $f_2 = 30.5846$, and $f_n > 2$ are unknown. Here $f_i$ differs from that of Narison ($f_i = 11/3$), but is in agreement with the equal mass case of Eq. (3.8) of Ref. [11]. Narison does not quote a value for $f_2$. Similarly, for the $D = 4,6$ contributions, one finds [11,8], with $\rho$ representing the deviation of the light 4-quark condensate from its vacuum saturation value,

$$\delta_{\nu}^{(4)} - \delta_{s}^{(4)} = \frac{36 \pi^2 a^3}{m_i^4} \left[ \langle m_i \bar{s}s \rangle - \langle m_i \bar{u}u \rangle \right]$$

$$+ \frac{36}{m_i^4} \left( \frac{m_i^2}{m_j^2} - 1 \right),$$

(8)

$$\delta_{\nu}^{(6)} - \delta_{s}^{(6)} = \frac{1792 \pi^4}{27} \rho a(m_i^2) \left[ \langle \bar{u}u \rangle^2 - \langle \bar{s}s \rangle^2 \right].$$

(9)

Note that the perturbative series in Eq. (7) converges much more slowly than does the analogous mass-independent perturbative contribution to inclusive $\tau$ decay. $P_{\text{incl}}(a) = 1 + a + 5.2023 a^2 + 26.366 a^3 + \cdots$. For $m_i = m_j$, e.g., using $\alpha(m_i^2) = 0.351 \pm 0.016$ [21], $P_{\text{incl}}(a(m_i^2)) = 1 + 0.1114 + 0.0645 + 0.0365 + \cdots$, while $P_{\text{max}}(a(m_i^2)) = 1 + 0.4828 + 0.3981 + \cdots$. Assumptions about the convergence of $P_{\text{max}}(a)$ based on analogy on the behavior of $P_{\text{incl}}(a)$ can, thus, not be expected to be reliable.

For the hadronic side of the sum rule, neglecting isospin breaking, one has, using the narrow width approximation for the isoscalar contributions [8],

$$R_{\nu}^{33} = \frac{3|V_{ud}|^2 S_{\text{EW}}}{2\pi \alpha_{\text{EM}}} \int_0^{m_i^2} ds \left( 1 - \frac{s}{m_i^2} \right)^2$$

$$\times \left[ 1 + \frac{2s}{m_i^2} \right] \frac{1}{m_i^2} \alpha_{\text{hadrons}}^{(s-1)} \left( \frac{\omega - \phi}{m_i^2} \right)^2,$$

(10)

$$R_{\nu}^{38} = \frac{18\pi |V_{ud}|^2 S_{\text{EW}}}{\alpha_{\text{EM}}}$$

$$\times \left[ 1 - \frac{m_i^2}{m_j^2} \right] \left( 1 + \frac{2m_i^2}{m_j^2} \right) \frac{m_i^4}{m_j^4} \frac{R_{\nu}^{38} - e^+ e^-}{m_i^2}$$

$$+ \left( \omega \rightarrow \phi \right) + \cdots,$$

(11)
where $+ \cdots$ refers to continuum and higher resonance contributions, which are small for $m_s$ less than $\sim 1.6$ GeV, and have been estimated by Narison [8]. Unfortunately, there turns out, numerically, to be a high degree of cancellation (to the 10–15% level) in $R_{33}^v - R_{38}^v$. With $m_t = 1.4, 1.6$ GeV, and Narison’s evaluation of the hadronic side, for example,

$$\left[ R_{33}^v - R_{38}^v \right]_{m_t=1.4\text{ GeV}} = (1.853 \pm 0.072)$$

$$- (1.581 \pm 0.066) = 0.272 \pm 0.098, \quad (12)$$

$$\left[ R_{33}^v - R_{38}^v \right]_{m_t=1.6\text{ GeV}} = (1.793 \pm 0.070)$$

$$- (1.626 \pm 0.069) = 0.167 \pm 0.098. \quad (13)$$

Isospin breaking at the few % level in the individual terms is, therefore, not necessarily negligible in the difference, especially if, as one would expect, e.g., for $\rho - \omega$ mixing effects, the signs of the effect were to be opposite in the two cases.

We now describe the input to the present version of analysis of the Narison sum rule. Modifications to the OPE side of the sum rule are minor. First, we use Narison’s updated value of the $D = 6$ 4-quark condensate, $\rho_\alpha \langle \bar{u}u \rangle^2 = (5.8 \pm 0.9) \times 10^{-4}$ GeV$^6$ [17], and, in evaluating the difference of light and strange quark terms, allow $\langle \bar{s}s \rangle / \langle \bar{u}u \rangle$ to vary between 0.7 and 1, as in Ref. [7]. Second, for the (very small) $D = 4$ terms, we take the light quark condensate to be given by the GMOR relation, and the strange quark condensate by the kaon version thereof, multiplied by a factor, $\bar{c}$, (varied between 0.5 and 1 to account for possible $SU(3)$ breaking). The ratio $m_s / \bar{m} = 24.4 \pm 1.5$, determined from ChPT to 1-loop [18], is also used. Third, since the 4-loop $\beta(a)$ [19] and $\gamma(a)$ [20] functions are now available, we employ throughout the 4-loop expressions for $a(Q^2)$ and $\bar{m}(Q^2)$, fixing the 3-flavor scale parameter, $\Lambda$, from the recent ALEPH $\tau$ decay data analysis [21].

The slow convergence noted above for $P_{\text{mass}}(a)$ at $m_t = m_s$ is even more accentuated at lower scales (e.g., for $m_t = 1.2, 1.4$ GeV, $P_{\text{mass}} = 1 + 0.672 + 0.736 + \cdots$ and $1 + 0.580 + 0.548 + \cdots$, respectively). An estimate of the $\mathcal{O}(a^3)$ term in $P_{\text{mass}}(a)$ thus appears crucial. We estimate $f_3$ using the procedure of Ref. [22] (CKS). For the three cases where the $\mathcal{O}(a^3)$ coefficient of quadratically-mass-dependent observables are known, the resulting estimates are accurate to $\pm 25$ [22]. The two possible versions, labelled FAC and PMS, yield $[f_3]_{\text{FAC}} = 288.0$ and $[f_3]_{\text{PMS}} = 290.1$. In contrast, [1,1] and [0,2] Padè estimates yield $[f_3]_{1,1} = 215.9$ and $[f_3]_{0,2} = 183.7$. To be conservative, we take $f_3 = 200 \pm 200$.

On the hadronic side the major change is that we now make isospin-breaking corrections to the extracted isovector and isoscalar spectral functions. These turn out to be quite large. Such corrections are unavoidable in the isoscalar case, where $e^+e^- \rightarrow$ hadrons is the only source of experimental data. In the isovector case, we will combine $\rho^+e^- \rightarrow$ hadrons and $\tau$ decay data in order to reduce the errors on $R_{33}^v$. For more details on the isospin breaking corrections to the $e^+e^- \rightarrow$ hadrons data, see Ref. [23].

The necessity of such corrections is obvious. Just as isospin breaking is observed through $\rho-\omega$ interference in $e^+e^- \rightarrow$ hadrons, so the vector meson electromagnetic (EM) decay constants, $F_{\rho,\omega}^\text{EM}$, will have both isospin-conserving and isospin-violating pieces. Defining $F_{\rho}^\prime$ by $\langle 0 | J_{\rho}^\prime | V, \lambda \rangle = F_{\rho}^\prime m_v e^\lambda$ for $a = 3, 8$, we see that $F_{\rho,\omega}^\text{EM} = F_{\rho}^\prime + \frac{1}{V} F_{\rho,\omega}^\text{V}, \quad F_{\rho,\omega}^\text{EM}$ vanish in the isospin limit, but will be non-zero in the real world. The $\rho$ contribution to the EM spectral function then contains an isospin-conserving piece proportional to $F_{\rho}^\text{EM}$ and an isospin-violating piece proportional to $F_{\rho}^\text{EM} F_{\omega}^\text{EM}$. The latter (associated with the 38 portion of $\rho_{\text{EM}}$) must be excluded in determining the $\rho$ contribution to $R_{38}^v$. Similarly, contributions proportional to $F_{\rho}^\text{EM} F_{\omega}^\text{EM}$ should be removed from the physical EM widths to obtain the $\omega$ and $\phi$ contributions to the 88 spectral function. The evaluation of the (unmeasured) isospin-breaking decay constants is accomplished by performing a QCD sum rule analysis of the mixed-isospin correlator $\langle 0 | T(J_{\rho}^\prime J_{\omega}^\prime) | 0 \rangle$ [23,25], for which the resonance contributions are directly proportional to the product $F_{\rho}^\text{EM} F_{\omega}^\text{EM}$. Combining the results for these products with the experimental values for $F_{\rho}^\text{EM}$, one extracts $F_{\rho}^\prime$ and $F_{\omega}^\prime$ separately. The sum rule provides good constraints on the product for the $\rho$ and $\omega$, weak constraints on the $\phi$, and is insensitive to higher resonance contributions. (Details of the analysis, in relation to CVC tests and the extraction of the sixth order ChPT low-energy constant, $Q$ [26,27], may be found in Ref. [23].) One finds $F_{\rho}^\text{EM} = 2.8 \pm 1.1$ MeV.
\( F^{(3)} = -4.2 \pm 1.5 \text{ MeV} \) and \( F^{(3)} = 0.21 \pm 0.21 \text{ MeV} \) (cf., \( F^{\text{EM}}_{\rho} = 154 \pm 3.6 \text{ MeV} \), \( F^{\text{EM}}_{\omega} = 45.9 \pm 0.8 \text{ MeV} \) and \( F^{\text{EM}}_{\phi} = -79.1 \pm 2.3 \text{ MeV} \)) [23]. These results satisfy several physical naturalness criteria [23]. The \( \rho \), \( \omega \) and \( \phi \) contributions to the 33 and 88 vector current spectral functions are given by \( \{ \rho^{\text{EM}}(q^2)\} = \{ F^{\text{EM}}_{\rho} \delta(x) \} \) where \( \delta(x) \), \( V = \rho, \omega, \phi \), in the narrow width approximation, is the usual \( \delta \) function, while \( \delta(x) \) is the corresponding \( \rho \) Breit-Wigner. The standard extractions, in contrast, are obtained by replacing \( F^{3} \) with \( F^{\text{EM}}_{\rho} \) and \( F^{8} \) with \( \sqrt{3} F^{\text{EM}}_{\omega} \) for \( V = \omega, \phi \). The corrections necessary to produce the true resonance contributions to the 33 and 88 spectral functions are thus

\[
\left( \frac{F^{(3)}_{\rho}}{F^{\text{EM}}_{\rho}} \right)^2 = 0.979 \pm 0.0086 ,
\]

\[
\left( \frac{F^{(8)}_{\omega}}{\sqrt{3} F^{\text{EM}}_{\omega}} \right)^2 = 1.189 \pm 0.065 ,
\]

\[
\left( \frac{F^{(8)}_{\phi}}{\sqrt{3} F^{\text{EM}}_{\phi}} \right)^2 = 1.0054 \pm 0.0054 . \tag{14}
\]

The size of the overestimate in the case of the 33 spectral function is still noticeably smaller than the \( \sim 5\% \) errors on the \( e^+e^- \rightarrow \text{hadrons} \) cross-sections in the resonance region. Note that the scale (\( \sim 1\% \)) of the isospin-breaking contribution to \( F^{\text{EM}}_{\rho} \) corresponds, as one might expect (since the assumption of mixing dominance corresponds to a leading chiral order approximation [23]), to what is obtained by assuming dominance by \( \rho-\omega \) mixing, and then evaluating this mixing following the updated analyses of \( e^+e^- \rightarrow \pi^+\pi^- \) discussed in Refs. [24]. Note also that, in this approximation, and neglecting both the \( \rho \) width and the difference of the \( \rho \) and \( \omega \) masses, the \( \rho \) and \( \omega \) correction terms would cancel in spectral integrals. The approximate cancellation of the corrections associated with the results of Eq. (14), when one forms a sum rule involving the sum of \( \rho \) and \( \omega \) contributions, is thus simply a manifestation of the dominant role of \( \rho-\omega \) mixing. Without understanding this point, it is easy to be misled into thinking that the scale of the individual resonance isospin-breaking decay constants is set by that of the isospin-breaking terms in the corresponding OPE representation. The latter, however, is associated with the sum, which involves a rather close cancellation, rather than the scale of the individual terms. While the \( 1\% \) correction in the case of the \( \rho \) is, as just explained, quite natural, the \( \omega \) correction to the 88 spectral function (\( \sim 19\% \)) might, in contrast, seem unnaturally large to some readers. It is, however, rather easy to see that such a large correction is actually expected in this case. As is well-known, in the limit that the vector meson nonet is ideally mixed, but the octet vector current matrix elements are otherwise given by \( \text{SU}(3)_r \), the EM decay constant of the \( I = 1 \) component of the \( \rho \), \( F^{\text{EM}}_{\rho} \), is 3 times that of the \( I = 0 \) component of the \( \omega \), \( F^{\text{EM}}_{\omega} \). Writing \( \rho = \rho_1 + \epsilon \omega_1 \) and \( \omega = \omega_1 - \epsilon \rho_1 \), (1 denoting the isospin pure states), with \( \epsilon \sim \sigma(\delta m) \), the physical EM decay constants become

\[
F^{\text{EM}}_{\rho} = F^{\text{EM}}_{\omega} = \frac{F^{\text{EM}}_{\rho}}{\sqrt{3} F^{\text{EM}}_{\omega}} \approx F^{\text{EM}}_{\rho} \left( 1 + \frac{\epsilon}{3} \right) ,
\]

\[
F^{\text{EM}}_{\omega} = F^{\text{EM}}_{\rho} + \epsilon F^{\text{EM}}_{\rho} = F^{\text{EM}}_{\rho} \left( 1 - 3 \epsilon \right) . \tag{15}
\]

The fractional correction for the \( \omega \) should thus be \( \sim 9 \) times that for the \( \rho \), and of opposite sign. Both features are present in the results of the sum rule analysis.

In reanalyzing the sum rule, we would, ideally, prefer to work with \( m_t \) near \( m_t \), where the perturbative series \( \gamma_{\text{mass}}(\alpha) \) is better behaved. Unfortunately, at such scales, the \( \omega' \) contribution to \( R^{88}_{\rho} \) becomes important. Since the sum rule employed to estimate the isospin-breaking decay constants is insensitive to the \( \rho-\omega' \) region, we are unable to correct the \( \omega' \) contribution for isospin breaking, and must thus work at scales below \( m_t \sim 1.6 \text{ GeV} \), where this contribution is small (\( < 2\% \) [8]). The slow convergence of \( F_{\text{mass}}(\alpha) \), similarly, forces us to values of \( m_t \) above \( \sim 1.4 \text{ GeV} \).

The corrected values of \( R^{33}_{\rho} \), \( R^{88}_{\rho} \), their difference, and the resulting values of \( m_t \), are given, as a function of \( m_t \), in Table 1. \( R^{33}_{\rho} \) is obtained by combining information from \( e^+e^- \rightarrow \text{hadrons} \) with that from \( \tau \) decay. In the latter case, we have used the recent ALEPH tabulation of the isovector vector spectral function [28], re-fitting the portion of the unfolded distribution from threshold up to \( s = (1.6 \text{ GeV})^2 \) relevant to our analysis. The results turn out to be almost identical to those obtained by direct
Table 1
Hadronic input and extracted values of $m_y$

<table>
<thead>
<tr>
<th>$m_y$ (GeV)</th>
<th>$R_{11}^{33}$</th>
<th>$R_{11}^{88}$</th>
<th>$R_{11}^{33} - R_{11}^{88}$</th>
<th>$m_y$ (MeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4</td>
<td>1.890 ± 0.098</td>
<td>1.706 ± 0.057</td>
<td>0.184 ± 0.113</td>
<td>123.1 ± 123</td>
</tr>
<tr>
<td>1.5</td>
<td>1.842 ± 0.063</td>
<td>1.737 ± 0.056</td>
<td>0.105 ± 0.084</td>
<td>104.4 ± 104</td>
</tr>
<tr>
<td>1.6</td>
<td>1.800 ± 0.048</td>
<td>1.713 ± 0.056</td>
<td>0.087 ± 0.074</td>
<td>110.4 ± 110</td>
</tr>
</tbody>
</table>

numerical integration of the unfolded ALEPH distribution. Although the errors associated with the extraction of the isovector spectral function from $\tau$ decay data are a factor of $\sim 2$ smaller than for $e^+e^- \rightarrow$ hadrons, one must bear in mind that there remain at present unknown $\sigma(\alpha_{EM})$ corrections to the relation between the spectral function extracted in $\tau$ decay and that appearing in $e^+e^- \rightarrow$ hadrons [29]. Marciano [29] has assigned an additional uncertainty of $\sim 3\%$ to account for these corrections. We have added this error in quadrature with that quoted by ALEPH. For the estimate based on $e^+e^- \rightarrow$ hadrons data, we employ Narison’s evaluation of the uncorrected hadronic integral for $R_{11}^{33}$, and PDG96 [30] values for the $\omega$ and $\phi$ EM widths to obtain the uncorrected version of $R_{11}^{88}$. Quoted values for $m_y$ in Table 1 are in the $\overline{MS}$ scheme, at scale $\mu = 1$ GeV, using the 4-loop running, and correspond to central values for all input. The errors quoted for $R_{11}^{88}$ reflect both those from the PDG96 partial widths and those from the uncertainties in the theoretical analysis of the mixed-isospin correlator sum rule. As expected, the isospin-breaking corrections have a very significant impact on $R_{11}^{33} - R_{11}^{88}$. The decrease in the value of the difference also magnifies the effect of the experimental errors. The errors quoted for $m_y$ correspond to those on the hadronic integrals only (the first to that on $R_{11}^{33}$, the second to that on $R_{11}^{88}$). Errors associated with uncertainties in the remaining inputs are as follows: (1) for $\bar{c}_x$: $\sim \pm 0.4$ MeV; (2) for the light-flavor four-quark condensate: $\sim \pm 4$ MeV; (3) for $\langle \Delta \rangle / \langle \bar{u}u \rangle$: $< \pm 17$ MeV; (4) for $f_3^3$: $< \pm 14$ MeV; (5) for $\Lambda_y$: $< \pm 3$ MeV.

We see from Table 1 that the corrections to $R_{11}^{88}$ significantly lower the hadronic side of the sum rule, and hence $m_y$. Taking an average of the three determinations, we obtain

$$\bar{m}_y(1\text{ GeV}) = 113^{+35}_{-54}^{+32} \pm 23\text{ MeV} \quad (16)$$

the first two sets of (asymmetric) errors being associated with the experimental input and the last set with uncertainties in the input on the OPE side. The result, while significantly below that of conventional sum rule analyses, is compatible with that of Colangelo et al. [7], and within errors, also the low values obtained in some recent lattice analyses [31,32]. To understand that the errors associated with the experimental input are still very significant, note that, if we use only the $\tau$ decay data to determine the isovector contributions, the central values of $m_y$ are changed to 148.5, 131.5 and 134.5 MeV, for $m_y = 1.4$, 1.5 and 1.6 GeV, respectively. We should also point out that the existing extraction of $m_y$ based on flavor-breaking in hadronic $\tau$ decay [33] is incorrect as a result of an error in the input coefficients of the perturbative series given in Ref. [16] for the mass-dependent $D = 2$ terms entering the strangeness-changing current contribution to these decays [34].

3. Summary

We have shown that isospin-breaking corrections significantly alter the value of $m_y$ extracted from existing $e^+e^- \rightarrow$ hadrons data, and that the resulting changes to the Narison sum rule for $m_y$ produce a central value $\bar{m}_y \sim 110$ MeV (at a scale 1 GeV) ($\sim 140$ MeV if one uses only $\tau$ decay data for the isovector input to the sum rule). The errors on this result, which are not insignificant, are currently dominated by the errors on the experimental input. Improved experimental data would, of course, significantly reduce the errors on the hadronic side of the sum rule. An improved treatment of both the hadronic and OPE sides of the mixed-isospin vector current sum rule is also highly desirable, both as a check of the stability of the results for, and a means of potentially reducing the errors on the determination.
of the isospin-breaking vector meson decay constants.

Acknowledgements

The author would like to thank S. Narison for bringing the contents of Ref. [8] to his attention, and to acknowledge both the ongoing support of the Natural Sciences and Engineering Research Council of Canada, and the hospitality of the Special Research Centre for the Subatomic Structure of Matter at the University of Adelaide, where this work was performed. Useful discussions with Andreas Höcker on the ALEPH spectral function analysis are also gratefully acknowledged.

References

[34] K. Maltman, in preparation.
Angular multiplicity fluctuations in hadronic Z decays and comparison to QCD models and analytical calculations

L3 Collaboration

M. Acciarri aa, O. Adriani P, M. Aguilar-Benitez z, S. Ahlen k, J. Alcaraz z,
G. Alemanni v, J. Allaby q, A. Aloisio ac, M.G. Alviggi ac, G. Ambrosi s,
H. Anderhub aw, V.P. Andreev al, T. Angelescu m, F. Anselmo i, A. Arefiev ab,
T. Azemoon c, T. Aziz j, P. Bagnaia ak, L. Bakasay ar, R.C. Ball c, S. Banerjee j,
Sw. Banerjee j, K. Banicz at, A. Barczyk aw, A. Barillere q, L. Barone ak,
P. Bartalini ah, A. Baschirotto aa, M. Basile t, R. Battiston ah, A. Bay v, F. Becattini p,
U. Becker o, F. Behner aw, J. Berdugo z, P. Berges o, B. Bertucci ah, B.L. Betev aw,
S. Bhattacharya j, M. Biasini q, A. Biland aw, G.M. Bilei ah, J.J. Blaising d,
S.C. Blyth ai, G.J. Bobbink b, R. Bock a, A. Böhm a, L. Boldizsár n, B. Borgia ak,
D. Bourilkov aw, M. Bourquin s, D. Boutigny d, S. Braccini s, J.G. Branson an,
V. Brigljevic aw, I.C. Brock ai, A. Buffini P, A. Buijs as, J.D. Burger o, W.J. Burger s,
J. Busenitz ar, X.D. Cai o, M. Campanelli aw, M. Capell o, G. Cara Romeo i,
G. Carlino ac, A.M. Cartacci p, J. Casaus z, G. Castellini p, F. Cavallari ak,
N. Cavallo ac, C. Cecchi s, M. Cerrada z, F. Cesaroni w, M. Chamizo z,
Y.H. Chang ay, U.K. Chaturvedi t, S.V. Chekanov ac, M. Chemarin y, A. Chen ay,
G. Chen g, G.M. Chen g, H.F. Chen t, H.S. Chen g, M. Chen o, G. Chiefari ac,
C.Y. Chien e, L. Cifarelli am, F. Cindolo i, C. Civinini p, I. Clare o,
R. Clare o, H.O. Cohn af, G. Coignet d, A.P. Colijn b, N. Colino z, S. Costantini h,
F. Cotorobai m, B. de la Cruz z, A. Csilling n, T.S. Dai o, R. D’Alessandro p,
R. de Asmundis ac, A. Degré d, K. Deiters au, P. Denes aj, F. DeNotaristefani ak,
D. DiBitonto ar, M. Diemoz ak, D. van Dierendonck b, F. Di Lodovico aw,
C. Dionisi ak, M. Dittmar aw, A. Dominguez an, A. Doria ac, M.T. Dova r l,
E. Drago ac, D. Duchesneau d, P. Dunker b, I. Duran ao, S. Dutta j, S. Easo ah,
Yu. Efremenko af, H. El Mamouni y, A. Engler ai, F.J. Eppling o, F.C. Erné b,
J.P. Ernenwein y, P. Extermann s, M. Fabre au, R. Faccini ak, S. Falciano ak,
A. Favara p, J. Fay y, O. Fedin al, M. Felcini aw, B. Fenyi ar, T. Ferguson ai,
F. Ferroni ak, H. Fesefeldt a, E. Fiandrini ah, J.H. Field s, F. Filthaut ai, P.H. Fisher o,
Abstract

Local multiplicity fluctuations in angular phase space intervals are studied using factorial moments measured in hadronic events at $\sqrt{s} = 91.2$ GeV, which were collected by the L3 detector at LEP. Parton shower Monte Carlo programs agree well with the data. On the other hand, first-order QCD calculations in the Double Leading Log Approximation and the Modified Leading Log Approximation are found to deviate significantly from the data. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

The analytical perturbative approach (APA) to QCD jet physics combines perturbative QCD calculations [1] with the principle of Local Parton Hadron Duality (LPHD), which relates parton distributions to those of hadrons. LPHD [2] assumes that if the parton cascade is evolved down to a sufficiently low scale, hadronic distributions are proportional to partonic ones. All non-perturbative effects are thus reduced to a normalisation constant.

This approach has been quite successful in describing inclusive quantities such as the single-particle scaled momentum spectrum, $\xi = \ln(1/x_p)$, and charged particle multiplicities in $e^+e^-$ data at LEP energies. However, less inclusive quantities have met with less success [3].

In this paper we study local fluctuations of the charged particle multiplicity, which provides a new test of APA applied to many-particle inclusive densi-
ties. Such fluctuations have been studied for many years in terms of a variety of phase space variables [4], but only recently has substantial progress been made in analytical QCD calculations of these observables [5–7].

We have recently investigated [8] local multiplicity fluctuations using bunching parameters [9], which showed directly that local fluctuations inside jets are multifractal, as is expected from QCD calculations [5–7]. In this paper we extend this study and present a quantitative comparison of first-order QCD calculations [5–7] with data from the L3 experiment at LEP using normalized factorial moments of orders q = 2, . . . , 5 in angular phase space intervals. The relative angle between particles has in the past proved to be sensitive to aspects of the QCD parton shower. For example, particle flow (the ‘‘string’’ effect) [10–12] and angular correlations such as the particle-particle correlation asymmetry (PPCA) [11,13] have demonstrated gluon interference in the parton shower.

An analysis similar to the present one [14] for q = 2, 3 found that calculations in the Double Leading Log Approximation (DLLA) [5] tended to underestimate the data if one used Λ = 0.1 − 0.2 GeV for the QCD dimensional scale. However, reasonable agreement was found using an effective Λ = 0.04 GeV, which is very small compared to QCD estimates [15].

2. Analytical calculations

QCD calculations [6,7] for the normalized factorial moments (NFMs) [16], Fq(Θ), have the following scaling behavior

\[ F_q(\Theta) = \frac{\langle n(n-1)\cdots(n-q+1) \rangle}{\langle n \rangle^q} \alpha \left( \frac{\Theta_0}{\Theta} \right)^{(1-D_q)q^{-1}}. \]  

where Θ0 is the half opening angle of a cone around the jet-axis, Θ is the angular half-width window of rings around the jet-axis centered at Θ0 (see Fig. 1), and n is the number of particles in these rings.

![Fig. 1. A schematic representation of the measurements of the local fluctuations in the polar angle around jet axis (D = 1).](image)

Brackets, ⟨⟩, around a quantity denote the average of that quantity over all events. Finally, Dq is the so-called Rényi dimension. The analytical QCD expectations for Dq are as follows [6,7]:

1. In the fixed-coupling regime, for moderately small angular bins,

\[ D_q = \gamma_0(Q) \frac{q+1}{q}. \]  

where \( \gamma_0(Q) = \sqrt{2} C_\alpha Q / \pi \) is the anomalous QCD dimension calculated at \( Q = E\Theta_0 \), \( E = \sqrt{s} / 2 \), s is the square of the center of mass energy, \( \alpha_s \) is the strong coupling constant, and \( C_\alpha = 3 \) is the gluon color factor (equal to the number of colors).

2. In the running-coupling regime, for small bins, the Rényi dimension becomes a function of the size of the angular ring (\( \alpha_s(Q) \) increases with decreasing \( \Theta \)).

It is useful to introduce a new scaling variable [7],

\[ z = \frac{\ln(\Theta_0/\Theta)}{\ln(\Theta_0 / \Lambda)}. \]  

The maximum possible phase space region (\( \Theta = \Theta_0 \)) corresponds to \( z = 0 \).

There are three approximate expressions derived in DLLA which will be tested:

a) According to [6], the \( D_q \) have the form

\[ D_q = \gamma_0(Q) \frac{q+1}{q} \left( 1 + \frac{q^2 + 1}{4q^2 z} \right). \]
b) Another approximation has been suggested \cite{7}:
\begin{equation}
D_q = 2 \gamma_0(Q) \frac{q + 1}{q} \left( 1 - \sqrt{1 - z} \right). 
\end{equation}

c) A result has also been obtained for the cumulant moments, which converge to factorial moments for high energies \cite{5}:
\begin{equation}
D_q = 2 \gamma_0(Q) \frac{q - w(q, z)}{z(q - 1)},
\end{equation}
\begin{equation}
w(q, z) = q \sqrt{1 - z} \left( 1 - \frac{\ln(1 - z)}{2q} \right).
\end{equation}

Furthermore, an estimate for \( D_q \) has been obtained in the Modified Leading Log Approximation (MLLA) \cite{6}. In this case, Eq. (4) remains valid except that \( \gamma_0(Q) \) is replaced by an effective \( \gamma_0^{\text{eff}}(Q) \) depending on \( q' \):
\begin{equation}
\gamma_0^{\text{eff}}(Q) = \frac{q - 1}{2(q + 1)} + \frac{1}{4}, \quad (7)
\end{equation}
where
\begin{equation}
b = \frac{11C_A}{3} - \frac{2n_f}{3}, \quad B = \frac{1}{b} \left[ \frac{11C_A}{3} + \frac{2n_f}{3C_A^2} \right],
\end{equation}
and \( n_f \) is the number of flavors.

For our comparison of the data with the theoretical calculations quoted above, we use the following parameters:
\[ n_f = 3, \quad \Lambda = 0.16 \text{ GeV}. \]
This value of \( n_f \) is chosen since even at high energies the production of heavy flavors will rarely happen in the jet and consequently its evolution is still dominated by the light flavors \cite{17}. The value of \( \Lambda \) chosen is that found in tuning the JETSET 7.4 matrix element program \cite{18} on L3 data \cite{19} and in our recent determination of \( \alpha_s(n_f) \) \cite{20}.

For the angle \( \Theta_0 \), we consider two possibilities: \( \Theta_0 = 25^\circ \) and \( 35^\circ \). The first value, suggested by authors of two of the calculations \cite{21}, is the same as used in the DELPHI analysis \cite{14}. The larger value of \( \Theta_0 \) allows a larger range of \( \Theta \) to be studied.

The effective coupling constant is evaluated at \( Q = E\Theta_0 \). For \( \Theta_0 = 25^\circ \), one obtains \( \alpha_s(E\Theta_0) = 0.144 \) according to the first-order QCD expression for \( \alpha_s(Q) \). This value leads to \( \gamma_0(E\Theta_0) = 0.525 \). For \( \Theta_0 = 35^\circ \), \( \alpha_s(E\Theta_0) = 0.135 \) and \( \gamma_0(E\Theta_0) = 0.508 \).

### 3. Experimental procedure

The analysis is based on data, corresponding to an integrated luminosity of 52 pb\(^{-1}\), collected by the L3 detector \cite{22} at a center of mass energy of \( \sqrt{s} = 91.2 \text{ GeV} \) during the 1994 LEP running period. Hadronic events are selected using information from the Central Tracking Detector (TEC) and the Silicon Microvertex Detector (SMD).

To obtain a sample with well-measured charged tracks, a selection is performed using tracks which have passed certain quality cuts. To ensure that the event lies within the full acceptance of the TEC and SMD, the direction of the thrust axis, as determined from the charged tracks, must satisfy \( |\cos \theta_{\text{thr}}| < 0.7 \). Events are then selected using the following criteria:
\[ \frac{\sum_i |p_i|}{\sqrt{\frac{\sum_i |p_i|}{\sum_i |p_{i\perp}|}}} > 0.15, \quad \frac{\sum_i |p_i|}{\sum_i |p_{i\perp}|} < 0.75, \quad \frac{\sum_i |p_{i\perp}|}{\sum_i |p_i|} < 0.75, \quad N_{\text{ch}} > 4, \]
where \( p_i \) is the momentum of particle \( i \) and the sum runs over all tracks of an event, and where \( N_{\text{ch}} \) is the number of charged tracks. The resulting sample contains about 1.0 million events.

In this paper we study fluctuations in small angular bins. For the grouping of tracks into these bins, the resolution of the angle between pairs of tracks is of crucial importance. For this reason we impose additional stringent quality cuts on track reconstruction, which results in rejection of 39% of the tracks. With this selection we achieve very good agreement between data and simulation for the distributions of the difference in angle between pairs of tracks for both the azimuthal angle about, and the polar angle with respect to, the beam \cite{8}. 

\[ M. \text{Acciarri et al.} / \text{Physics Letters B} 428 (1998) 186--196 191 \]
An NFM calculated from the data is corrected for detector effects by a correction factor determined from two Monte Carlo samples. Events generated with the JETSET 7.4 parton shower (ps) program [18] including initial-state photon radiation are passed through a full detector simulation [23] including time-dependent variations of the detector response based on continuous detector monitoring and calibration. It has been reconstructed with the same program as the data and passed through the same selection procedure. The resulting sample is referred to as detector level MC. Another sample, called generator level MC, is generated directly from JETSET. It contains all charged final-state particles with a lifetime $c\tau > 1$ cm and is generated without initial-state photon radiation, Dalitz pairs or Bose-Einstein correlations, since these effects are not included in the analytical QCD calculations.

From these two samples a correction factor is found: $C_q = F^{gen}_q / F^{det}_q$, where $F^{gen}_q$ and $F^{det}_q$ are the values of the NFM of order $q$ calculated from the generator level and detector level, respectively. The corrected NFM is then given by $F_q = C_q F^{raw}_q$, where $F^{raw}_q$ is the NFM calculated directly from the data. The correction is of the order of 3% for $F_2$, increasing to approximately 5% for $F_3$.

The resolution of the L3 detector for a number of relevant variables has been estimated [24]. The resolution of polar angle defined with respect to the thrust axis is found to be approximately 0.01 radians. For higher orders NFMs, the minimum angle $Q$ used in this study is chosen according to the many-particle resolutions studied in [24].

The errors on the results include both statistical and systematic errors on the raw quantities and on the correction factors. The systematic errors on the raw quantities, found from variation of track quality cuts and event selection cuts were found to be negligible. The systematic error on the correction factors is taken as half of the difference between the correction factors determined using JETSET and those using HERWIG 5.9 [25].

4. Results

The sphericity axis is used to define the jet axis. To increase statistics, we evaluated the NFMs in each sphericity hemisphere of an event and averaged the results, thus assuming that the local fluctuations in each hemisphere are independent. Fig. 2 shows the experimental results on the behavior of the NFMs as a function of the scaling variable $z$, for $\Theta_0 = 25^\circ$ and $\Lambda = 0.16$ GeV, compared to Monte Carlo model predictions on the partonic and hadronic levels. The shaded areas in this figure, and in Fig. 3 Fig. 4, represent the statistical errors on the model predictions.

4.1. Comparison with Monte Carlo models

The data in Fig. 2 are compared with the predictions of the JETSET, HERWIG, and ARIADNE [26] parton shower models at both the hadronic and partonic levels. All three models have been tuned to reproduce global event-shape and single-particle inclusive distributions [27,28]. The hadronic-level predictions of the models give a good description of the fluctuations. The effect of heavy flavors (c and b quarks) has been estimated by rejecting these flavors in JETSET. The effect was found to be negligible.

---

The Bose-Einstein modelling of JETSET is used in ARIADNE; HERWIG contains no Bose-Einstein model; and JETSET was used with its Bose-Einstein modelling turned off.
The data and the hadronic level of the models saturate later than does the partonic level.

It is expected [7] that hadronization effects would largely cancel in the ratio \( F_q(z)/F_q(0) \). In addition, this ratio eliminates a theoretical ambiguity in the normalization of the NFMIs, i.e., in \( F_q(0) \). In terms of \( F_q(z)/F_q(0) \), the power law of Eq. (1) can be rewritten as

\[
\ln \frac{F_q(z)}{F_q(0)} = z(1-D_q)(q-1)\ln \frac{E\theta_0}{\Lambda}.
\]

(8)

The behavior of \( \ln(F_q(z)/F_q(0)) \) as a function of \( z \) is shown in Figs. 3 and 4 for the partonic and hadronic levels, respectively. The partonic level predictions of the models are indeed much closer to the data, and the differences between partonic and hadronic levels are decreased, particularly for the higher-order moments. Here too the hadronic level of the models provides a satisfactory description of the data. The degree of similarity between partonic and hadronic level MC predictions can be interpreted as a measure of the degree of validity of LPHD. We note that there is a greater difference between the partonic and hadronic levels of HERWIG than of JETSET with ARIADNE lying in between. We also note that the average number of partons is about 8.6.

Fig. 3. \( F_q(z)/F_q(0) \) as a function of the scaling variable \( z \), for \( \theta_0 = 25^\circ \) and \( \Lambda = 0.16 \text{ GeV} \), compared to Monte Carlo model predictions on the partonic level.

Fig. 4. \( F_q(z)/F_q(0) \) as a function of the scaling variable \( z \), for \( \theta_0 = 25^\circ \) and \( \Lambda = 0.16 \text{ GeV} \), compared to Monte Carlo model predictions on the hadronic level.

10.0, and 10.2 for HERWIG, JETSET, and ARIADNE, respectively.

4.2. Comparison with analytical calculations

The comparison of the analytical QCD calculations (Eqs. (2), (4)–(7)) with the corrected data is shown in Fig. 5 for \( \Lambda = 0.16 \text{ GeV} \) and \( \theta_0 = 25^\circ \). For the second order moment, running \( \alpha_s \) calculations lead to the saturation effects observed in the data, but significantly underestimate the observed signal. Predictions for the higher moments are too low for low values of \( z \), but tend to overestimate the data at larger \( z \). The fixed coupling regime (thin solid lines) approximates the running coupling regime for small \( z \), but does not exhibit the saturation effect seen in the data. The DLLA approximations (Eqs. (4)–(6)) differ significantly at large \( z \), with the calculations from cumulants (Eq. (6)) showing the strongest saturation effect. The MLLA predictions are rather similar to the DLLA results of Eq. (4).

We have also compared (not shown) the data and the QCD predictions for \( \theta_0 = 35^\circ \). Both the data and the predictions rise more rapidly than for \( \theta_0 = 25^\circ \). This indicates that fluctuations are larger for phase space regions containing a larger contribution from hard gluon radiation. However, the disagreement
between data and predictions is similar to that for \( \Theta = 25^\circ \).

In a study using the parameterization of [5] (Eq. (6)), DELPHI has found better agreement with their data by decreasing the value of \( \Lambda \) to 0.04 GeV [14]. A smaller effective value makes the coupling constant smaller, which expands the range of validity of the perturbative calculations (for \( \Lambda = 0.04 \) GeV, \( \alpha_s(E \Theta) = 0.112, \gamma_0(E \Theta) = 0.46 \)). Fig. 6 shows the case of \( \Lambda = 0.04 \) GeV for our data. While the agreement for small \( z \) is indeed better, it becomes worse for \( z > 0.3 \), where contributions from higher-order perturbative QCD and hadronization are expected to be larger. We have varied \( \Lambda \) in the range of 0.04 – 0.25 GeV and found that there is no value of \( \Lambda \) in this range which produces agreement for all orders of NFMs. Increasing the number of active flavors, \( n_f \), to 4 or 5 leads to worse agreement.

4.3. Discussion

The first-order calculations of the DLLA and MLLA of perturbative QCD are shown to be in disagreement with the local fluctuations observed in hadronic Z decay. This occurs both for standard values of \( \Lambda \) (\( \Lambda = 0.16 \) GeV) and for small values (\( \Lambda = 0.04 \) GeV). In the latter case, a reasonable estimate for \( z \sim 0.3 \) can be obtained, consistent with the DELPHI conclusion [14]. However, in this case, the theoretical NFMs strongly overestimate the data for relatively large \( z \) (small \( \Theta \)), where contributions from higher-order perturbative QCD are larger.

On the other hand, the MC models all agree well with the data.

Likely reasons for the failure of the calculations are their asymptotic character, which corresponds to an infinite number of partons in an event, and their lack of energy-momentum conservation, features which are taken into account in the MC models. A similar conclusion was reached from a comparison of factorial and cumulant moments in quark and gluon jets [29]. Further, a recent theoretical study [30] of energy conservation in triple-parton vertices shows that the energy conservation constraint is indeed sizeable and leads to a stronger saturation effect. Note that the MLLA predictions used here are not from a full MLLA calculation. This MLLA calculation only modifies \( \gamma_0 \), while retaining the DLLA parameterization of the \( z \)-dependence of the NFMs, which is only asymptotically correct.

Another contribution to the failure of the predictions can lie with the local parton-hadron duality...
hypothesis, which is used to justify comparison of the analytical QCD calculations with hadronic data. In JETSET the difference between parton- and particle-level predictions are large for $F(x)/F(0)$ but small for $F_q(z)/F_q(0)$. However, for HERWIG and ARIDNE this is not the case. At the shower cut-off scales of these models, the hadronization effects are thus still important and depend on $z$. Thus LPHD does not apply at these cut-off scales. It is conceivable that lowering the cut-off in these models below the current value of about 1 GeV would result in smaller hadronization effects and better agreement. However, we consider it unlikely that such a model could successfully describe other aspects of the data, such as production rates for baryons and high-mass meson resonances.

5. Conclusions

Monte Carlo models incorporating a coherent parton shower agree well with the data. On the other hand, first-order calculations in the DLLA and MLLA of perturbative QCD disagree with the local fluctuations observed in hadronic Z decay. The asymptotic nature of the calculations and their inadequate treatment of energy-momentum conservation appear to be the most likely reasons for the failure of the calculations. The Monte Carlo results at parton level indicate that the influence of hadronization is not in agreement with the LPHD assumption.

Acknowledgements

We wish to express our gratitude to the CERN accelerator divisions for the excellent performance of the LEP machine. We acknowledge the effort of all the engineers and technicians who have participated in the construction and maintenance of this experiment. We thank B. Buschbeck, M. Chmeissani, I. Dremin, V.I. Kuvshinov, P. Lipa, F. Mandl, J.-L. Meunier, P. Nason, W. Ochs, R. Peschanski, T. Sjöstrand and B. Webber for helpful discussions.

References


[23] The L3 detector simulation is based on GEANT3, see R. Brun et al., CERN report CERN DD/EE/84-1 (Revised), 1987, and uses GHEISHA to simulate hadronic interactions, see H. Fesefeldt, RWTH Aachen report PITHA 85/02, 1985.


[29] OPAL Collaboration, K. Ackerstaff et al., CERN-PPE/97-105 Submitted to Z. Phys. C.

Search for a new gauge boson in $\pi^0$ decays

NOMAD Collaboration

D. Verkindt \textsuperscript{b}, J-M. Vieira \textsuperscript{j}, T. Vinogradova \textsuperscript{k}, M-K. Vo \textsuperscript{s}, S.A. Volkov \textsuperscript{m}, F. Weber \textsuperscript{d}, T. Weisse \textsuperscript{i}, M. Werlen \textsuperscript{j}, F. Wilson \textsuperscript{i}, L.J. Winton \textsuperscript{l}, B.D. Yabsley \textsuperscript{t}, H. Zaccone \textsuperscript{s}, K. Zuber \textsuperscript{f}

\textsuperscript{a} Univ. of Massachusetts, Amherst, MA, USA
\textsuperscript{b} LAPP, Annecy, France
\textsuperscript{c} Johns Hopkins Univ., Baltimore, MD, USA
\textsuperscript{d} Harvard Univ., Cambridge, MA, USA
\textsuperscript{e} Univ. of Calabria and INFN, Cosenza, Italy
\textsuperscript{f} Dortmund Univ., Dortmund, Germany
\textsuperscript{g} JINR, Dubna, Russia
\textsuperscript{h} Univ. of Florence and INFN, Florence, Italy
\textsuperscript{i} CERN, Geneva, Switzerland
\textsuperscript{j} University of Lausanne, Lausanne, Switzerland
\textsuperscript{k} UCLA, Los Angeles, CA, USA
\textsuperscript{l} University of Melbourne, Melbourne, Australia
\textsuperscript{m} Institute for Nuclear Research, INR Moscow, Russia
\textsuperscript{n} Univ. of Padova and INFN, Padova, Italy
\textsuperscript{o} LPNHE, Univ. of Paris, Paris VI and VII, France
\textsuperscript{p} Univ. of Pavia and INFN, Pavia, Italy
\textsuperscript{q} Univ. of Pisa and INFN, Pisa, Italy
\textsuperscript{r} Roma-III Univ., Rome, Italy
\textsuperscript{s} DAPNIA, CEA Saclay, France
\textsuperscript{t} Univ. of Sydney, Sydney, Australia
\textsuperscript{u} ANSTO, Sydney, Menai, Australia
\textsuperscript{v} Univ. of Urbino, Urbino, and INFN Florence, Italy
\textsuperscript{w} Rudjer Bošković Institute, Zagreb, Croatia

Received 30 March 1998

Editor: L. Montanet

Abstract

A search was made for a new light gauge boson $X$ which might be produced in $\pi^0 \rightarrow \gamma + X$ decay from neutral pions generated by 450 GeV protons in the CERN SPS neutrino target. The $X$’s would penetrate the downstream shielding and be observed in the NOMAD detector via the Primakoff effect, in the process of $X \rightarrow p^0$ conversion in the external Coulomb field of a nucleus. With $1.45 \times 10^{18}$ protons on target, 20 candidate events with energy between 8 and 140 GeV were found from the analysis of neutrino data. This number is in agreement with the expectation of $18.1 \pm 2.8$ background events from standard neutrino processes. A new 90 \% C.L. upper limit on the branching ratio $Br(\pi^0 \rightarrow \gamma + X)$ is obtained.

1. Introduction

Many extensions of the Standard Model such as GUTs \cite{1}, super-symmetric \cite{2}, super-string models \cite{3} and models including a new long-range interaction, i.e. the fifth force \cite{4}, predict an extra $U(1)$ factor and therefore the existence of a new gauge boson $X$ corresponding to this new group. The predictions for the mass of the $X$ boson are not very firm and it could be light enough ($M_X \ll M_Z$) for searches at low energies.

If the mass $M_X$ is of the order of the pion mass, an effective search could be conducted for this new vector boson in the radiative decays of neutral pseudoscalar mesons $P \rightarrow \gamma + X$, where $P = \pi^0, \eta$, or $\eta'$, because the decay rate of $P \rightarrow \gamma + \ldots$
any new particles with spin 0 or $\frac{1}{2}$ proves to be negligibly small [5]. Therefore, a positive result in the direct search for these decay modes could be interpreted unambiguously as the discovery of a new light spin 1 particle, in contrast with other experiments searching for light weakly interacting particles in rare K, $\pi$ or $\mu$ decays [5,6].

From the analysis of the data from earlier experiments, constraints on the branching ratio for the decay of $P \rightarrow \gamma + X$ range from $10^{-7}$ to $10^{-3}$ depending on whether $X$ interacts with both quarks and leptons or only with quarks. Since in the first case $X$ is a short lived particle decaying mainly to $e^+e^-$ or $\nu\bar{\nu}$ pairs, we will only consider the second case, where $X$ is a relatively long-lived particle [6].

Direct searches for a signal from $\pi^0 \rightarrow \gamma + X$ decay have been performed in a few experiments with two different methods: i) searching for a peak in inclusive photon spectra from two-body $\pi^0 \rightarrow \gamma + \text{nothing}$ decays, where ‘‘nothing’’ means that $X$ is not detected because it either has a long life time or decays into $\nu\bar{\nu}$ pairs (Refs. [7–9]); and ii) searching for a peak in the invariant mass spectrum of $e^+e^-$ pairs from $\pi^0$ decays, which corresponds to the decay $X \rightarrow e^+e^-$ [10].

The best experimental limit on the decay $\pi^0 \rightarrow \gamma + X$ was obtained recently by the Crystal Barrel Collaboration at CERN [9]. Using $pp$ annihilations as a source of neutral pions, they searched for a single peak in the inclusive photon energy spectrum. The branching ratio limit of $(6$ to $3) \times 10^{-5}$ (90% C.L.) was obtained for $65 < M_X < 125$ MeV. A less stringent upper limit $BR(\pi^0 \rightarrow \gamma + X) < (3$ to $0.6) \times 10^{-4}$ was obtained for the mass region $0 < M_X < 65$ MeV/$c^2$. This result is valid for the case where $X$ is a long-lived particle or it decays mainly to $\nu\bar{\nu}$ pairs.

In this paper, we present a more sensitive upper limit on the branching ratio of the decay $\pi^0 \rightarrow \gamma + X$ obtained by using a new method [11] from the analysis of high energy neutrino data taken by the NOMAD experiment at CERN.

2. The NOMAD detector

The NOMAD detector, designed to search for a neutrino oscillation signal in the CERN SPS wide-band neutrino beam, is described in detail in Ref. [12]. A sketch of the NOMAD detector is shown in Fig. 1. It consists of a number of subdetectors most of which are located inside a 0.4 T dipole magnet with a volume of $7.5 \times 3.5 \times 3.5$ m³. The relevant features for the present study will be briefly mentioned.

The complete active target consists of 44 drift chambers (DC) mounted in 11 modules [13]. The target is followed by a transition radiation detector (TRD) to enhance $e/\pi$ separation and by a lead-glass electromagnetic calorimeter (ECAL) with an upstream preshower detector (PRS). Five additional drift chambers are installed in the TRD region.

Each of the 9 TRD modules [14] consists of a radiator followed by a detection plane of vertical straw tubes.

The ECAL consists of 875 lead-glass Cerenkov counters of TF1-000 type arranged in a matrix of 35 rows by 25 columns. Each counter is about 19 radiation lengths deep. The energy resolution is $\sigma/E = 0.01 + 0.032/\sqrt{E}$, where $E$ is the shower energy in GeV. A more detailed description of the ECAL is given elsewhere [15].

The PRS is composed of two planes of proportional tubes (286 horizontal and 288 vertical tubes) preceded by a 9 mm thick (1.6 $X_0$) lead plate, providing a spatial resolution $\sigma_{x,y} \approx 1$ cm/$\sqrt{E}$, where $E$ is the shower energy in GeV. The fiducial mass of this detector is about 700 kg.

The trigger for neutrino interactions in the target is provided by two planes of scintillation counters $T_1$ and $T_2$. A veto in front of the magnet rejects upstream neutrino interactions and muons incident on the detector. Neutrino interactions in the PRS or ECAL are collected by a $T_1 \times T_2 \times ECAL$ trigger specially designed for this purpose. The ECAL signal is obtained as the OR of all counter signals exceeding a threshold of about 0.8 GeV. The timing of this signal depends on the deposited energy. For energies above 3 GeV a time resolution of a few ns and a trigger efficiency of 100% are obtained. The average rate of the $T_1 \times T_2 \times ECAL$ trigger is about $3 \times 10^{13}$ protons on the neutrino target (p.o.t.).

A hadronic calorimeter(HCAL) and a set of 10 drift chambers placed behind the magnet provide an estimate of the energy of the hadronic component in the event and muon identification, respectively. The
HCAL is an iron-scintillator sampling calorimeter consisting of 11 iron plates, 4.9 cm thick, separated by 1.8 cm gaps in which scintillator paddles 3.6 m long 1 cm thick, and 18.3 cm wide are installed. The HCAL active area is 3.6 m wide by 3.5 m high, and approximately 3.1 hadronic interaction lengths deep.

3. Method of search

If the decay $\pi^0 \rightarrow \gamma + X$ exists, one expects a flux of high energy $X$ bosons from the SPS neutrino target, since $\pi^0$'s are abundantly produced in the forward direction by 450 GeV protons either in the beryllium target or in the beam dump following the decay tunnel. If $X$ is a long-lived particle, this flux would penetrate the downstream shielding without significant attenuation and would be observed in the NOMAD detector via the Primakoff effect, namely in the conversion process $X \rightarrow \pi^0$ in the external Coulomb field of a nucleus [11] (see Fig. 2).

Because the cross section for $X \rightarrow \pi^0$ conversion is proportional to $Z^2$, we searched for these events in the lead of the preshower detector. The experimental signature of $X \rightarrow \pi^0$ conversion is a single high energy $\pi^0$ decaying into two photons which results in a single isolated electromagnetic shower in the ECAL. The corresponding ECAL cluster should be matched to the PRS cluster from the converted photons and must not be accompanied by a significant activity in any of the other NOMAD subdetectors.
The occurrence of $X \rightarrow \pi^0$ conversion would appear as an excess of neutrino-like interactions in the PRS with pure electromagnetic final states above those expected from Monte Carlo predictions of standard neutrino interactions. The expected energy spectrum of $X$-bosons at the NOMAD detector is shown in Fig. 3 for a mass $M_X = 10$ MeV. The energy spectra of $\pi^0$'s produced in the neutrino target and in the beam dump have been obtained with the same detailed GEANT 16 simulation used to predict the neutrino flux distributions at the NOMAD detector. The $X$-spectrum is considerably harder than that of neutrinos which is also shown for comparison.

4. Data sample and event selection

This analysis is based on the data taken during the first half of the neutrino run in 1995, in which the NOMAD target was only partially installed and consisted of four drift chamber modules placed upstream of the TRD detector. The integrated number of protons delivered to the neutrino target during this period was about $1.45 \times 10^{18}$.

Monte Carlo simulations use the LEPTO 6.1 event generator [17] supplemented by event generators for resonance, quasi-elastic and coherent neutrino processes and a full detector simulation based on GEANT [16]. These simulations together with direct measurements of subdetector occupancy during the neutrino spill with a random trigger [12] show that $\nu$ and $X \rightarrow \pi^0$ events in the PRS are accompanied by no significant activity in the DC and TRD. For this reason a simple cut on the number of hits $N_{hit}^{DC}$, $N_{hit}^{DC(\text{TRD})}$ and $N_{hit}^{\text{TRD}}$ in the DC, in the DC located in the TRD region, and in the TRD, respectively, were used as a veto to suppress events from neutrino interactions which occurred in these regions or in the magnet coils. However, neutrino events, mainly with a neutral final state, which occurred in the near upstream PRS region can also pass these veto cuts and can be taken as neutrino events in the PRS. For this reason the upstream region was also included in the simulations and surviving events were considered as neutrino events in the PRS.

The selection criteria for $X \rightarrow \pi^0$ events are based on a full Monte Carlo simulation of $X \rightarrow \pi^0$ conversions in the NOMAD detector and rely on their properties as mentioned in Section 3, namely:

- the quality of the match between the ECAL shower and the corresponding PRS cluster;
- the total energy and the shower shape in the ECAL;
- the amount of energy deposition in the HCAL.

Candidate events were identified by the following simultaneous requirements:

- $DC \times TRD$ (no activity in the DC or TRD): $N_{hit}^{DC} \leq 10$, $N_{hit}^{DC(\text{TRD})} \leq 4$, $N_{hit}^{\text{TRD}} \leq 2$.
- $PRS \times ECAL > 8$ GeV): isolated PRS cluster matched to isolated ECAL cluster in both X and Y planes.

At most two ECAL clusters in the ECAL were allowed. The energy of the most energetic cluster had to be $E_{ECAL} > 8$ GeV and the second cluster had to be less than 0.3 GeV in order to reject pile-up events and ECAL noise. The shape of the most energetic cluster was fitted to the shape expected from an electromagnetic shower and the $\chi^2$ of the fit was required to be less than 20 [18]. The differences $\Delta X$ and $\Delta Y$ between the X and Y coordinates of the cluster in the PRS and the corresponding cluster in the ECAL were required to be $|\Delta X| < 2$ cm, $|\Delta Y| < 1.5$ cm. These conditions were used to identify isolated electromagnetic showers in the ECAL that origi-
nated from the conversion of photons from $\pi^0$'s produced in the preshower.

- **HCAL** (no HCAL activity): the total visible energy was required to be less than the HCAL noise threshold $E_{\text{HCAL}} < 0.4$ GeV.

This cut serves to identify electromagnetic energy in the ECAL. The reliability of the HCAL veto to suppress conventional neutrino events with an energy leakage to the HCAL detector was checked with straight through muons and with a selected sample of $\nu_\mu$ charged current (CC) events in the PRS.

- **MUON**: no track(s) in the muon chambers matched to the PRS cluster.

After applying these cuts to the initial sample of $4.83 \times 10^5$ events recorded with the ECAL trigger we found 20 candidate events for $X \rightarrow \pi^0$ conversion. The amount of background in this sample from standard $\nu_e$ and $\nu_\mu$ interactions was evaluated using the Monte Carlo (see Section 5).

Applying the $DC \times TRD \times PRS \times ECAL(>1$ GeV) $\times MUON$ selection criteria to the same initial sample we find 3691 events from $\nu_\mu$ CC interactions in the PRS. Here, $MUON$ denotes single muon tracks extrapolated back from the muon chambers to the PRS and matched to the PRS cluster. The spectra of energy deposited in the ECAL for these selected $\nu_\mu$ CC events and for simulated events which pass the same reconstruction program and selection cuts are found to be in agreement. The $T_1 \times T_2 \times ECAL$ trigger efficiency was obtained from a Monte Carlo simulation and was found to be 71% and 97% for $\nu_\mu$ CC and $X \rightarrow \pi^0$ events, respectively. Taking the overall selection efficiency into account, the total number of $\nu_\mu$ CC events interacting in the PRS was found to be $N_{\nu_\mu,\text{PRS}} = 2.29 \times 10^4$. The number of $\nu_\mu$ CC events was converted to the number of neutrons on target $N_{\nu_\mu}$ using the simulation of the neutrino flux at the NOMAD detector and the known cross section value for $\nu_\mu$ CC interactions, giving $N_{\text{pot}} = 1.36 \times 10^{18}$. This number agrees within 10% with the value of $N_{\text{pot}}$ measured by the beam monitors and was used further for normalisation. The main uncertainty results from the contribution of events from neutrino interactions in the downstream TRD region which pass the DC and TRD cuts and from backscattering in events occurring in the PRS. By varying the $DC$ and $TRD$ cuts, it was found that the systematic error in the number of $\nu_\mu$ CC events in the PRS resulting from backscattering is of the order of 10%.

5. Background events

The main background to $X \rightarrow \pi^0$ conversions is expected from the following neutrino processes occurring in the PRS or upstream PRS region with a significant electromagnetic component in the final state and with no significant energy deposition in the HCAL:

- $\nu_\mu$ CC interactions classified as muonless because the muon was not detected;
- inclusive $\pi^0$ production from $\nu_\mu$ neutral current (NC) interactions;
- coherent and diffractive $\pi^0$ production;
- quasi-elastic $\nu_\mu$ scattering.

- $\nu_\mu$ CC interactions;
- $\nu_e + Pb \rightarrow \nu_e + e^+ + e^- + Pb$ and from pile-up events were found to be negligible.

To evaluate the amount of background in the data sample, simulated events were processed through the same reconstruction program and selection criteria that were used for real neutrino data. Then, all background distributions were summed up, taking into account the corresponding normalisation factors. These factors were calculated from beam composition and cross sections of the well-known processes listed above. The total number of events interacting in the PRS and the number of expected candidate events after applying the selection criteria are given in Table 1 for each background process.

The total background in the data sample was estimated to be $18.1 \pm 2.8$ events, where statistical

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Overall background estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>item</td>
<td>$\nu_e$ CC</td>
</tr>
<tr>
<td>Total number of interactions in the PRS</td>
<td>22886</td>
</tr>
<tr>
<td>Number of expected candidate events</td>
<td>1.7 ± 1.0</td>
</tr>
</tbody>
</table>
and systematic errors are added in quadrature. The fraction of neutrino interactions in the PRS which satisfy all cuts is \(6 \times 10^{-4}\).

6. Result

Fig. 4 shows the combined background and candidate event energy spectra in the ECAL. The agreement between data and Monte Carlo is good. The overall efficiency for \(X \rightarrow \pi^0\) conversion detection was found to be \(\approx 24\%\). The inefficiency is mostly due to the requirement that at least one photon from \(\pi^0\) decay convert in the PRS.

By subtracting the number of expected background events from the number of candidate events we obtain \(N_{\chi^0} = 1.9 \pm 5.3\) events showing no excess of \(X \rightarrow \pi^0\) conversion-like events, and, hence, no indication for the existence of this process. The 90% C.L. upper limit for the branching ratio \(Br(\pi^0 \rightarrow \gamma + X)\) was calculated by using the following relation:

\[
N_{\chi^0}^{90\%} > \left[ Br(\pi^0 \rightarrow \gamma + X) \right]^2 \times \int \frac{d\phi(M_X,E_X,N_{pot})}{dE_X} \sigma_\gamma(M_X,E_X) \times dE_X \times \varepsilon_{sel} \times \rho \frac{N_A}{\Lambda} (1)
\]

where \(N_{\chi^0}^{90\%} = 8.7\) events is the 90% C.L. upper limit for the expected number of signal events, \(\varepsilon_{sel}\) is the selection efficiency, which was found to be practically independent of \(E_X\), \(f_\phi(M_X,E_X,N_{pot})\) and \(\sigma_\gamma(M_X,E_X)\) are the flux of \(X\) bosons for the given number of protons \(N_{pot}\) on the neutrino target and the cross section for \(X \rightarrow \pi^0\) conversion on lead, respectively, calculated for \(Br(\pi^0 \rightarrow \gamma + X) = 1\). In Eq. (1), \(M_X\) and \(E_X\) are the mass and energy of the \(X\) boson, respectively. The cross section \(\sigma_\gamma(M_X,E_X)\) is given in Ref. [11]. The total \(X\) flux per proton on target calculated as a function of the \(X\) boson mass is shown in Fig. 5. We note that \(Br(\pi^0 \rightarrow \gamma + X)\) appears twice in the formula for \(N_{\chi^0}^{90\%}\), through
the $X$ boson flux from the target and through the Primakoff mechanism.

The 90% C.L. branching ratio limit curve is shown in Fig. 6 together with the result of Ref. [9]. For the mass region $0 < M_X < 120$ MeV/c$^2$ the limit is

$$Br(\pi^0 \rightarrow \gamma + X) < (3.3 \text{ to } 1.9) \times 10^{-5}$$

Varying the cut on the ECAL energy deposition in the range from 5 to 50 GeV does not change the limit substantially, while above 50 GeV it becomes worse. The limit is valid for an $X$ boson lifetime $\tau_X > 10^{-9} M_X [\text{MeV/c}^2] s$. For the mass region $0 < M_X < 60$ MeV/c$^2$, the limit is approximately a factor 10 to 5 better than the best previously published limit obtained by the Crystal Barrel collaboration [9].

Our result can also be used to constrain the magnitude of the coupling of a hypothetical gauge $X$ boson to quarks [6,11]

$$g^2 < 2 \times 10^{-7} \left(1 - \frac{m_X^2}{m_q^2}\right)^{-3}$$

The attenuation of the $X$-flux due to $X$ interactions with matter was found to be negligible, since for $Br(\pi^0 \rightarrow XY) \lesssim 10^{-4}$ the $X$ boson mean free

path in iron is $\geq 300$ km, as compared with the iron and earth shielding total length of 0.4 km used in our beam. Furthermore the decay length for $E_X > 10$ GeV was estimated to be longer than $= 10$ km [11]. We note that limits on the branching ratio $Br(\eta, \eta' \rightarrow \gamma + X)$ could also be obtained from our data if the cross-sections for $\eta, \eta'$ productions in the forward direction in $p-Be$ collisions at 450 GeV were known.

Acknowledgements

We thank the SPS staff, the neutrino beam group and the technical staffs of the participating institutions for their vital contributions to the success of this experiment.

This experiment was supported by the Australian Research Council and the Department of Industry, Science and Technology, the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie of Germany, CERN, the Institut National de Physique Nucléaire et des Particules and the Commissariat à l’Energie Atomique of France, the Istituto Nazionale di Fisica Nucleare of Italy, the Institute for Nuclear Research of the Russian Academy of Sciences and Russian Foundation for Fundamental Research, the Fonds National Suisse de la Recherche Scientifique, the U.S. Department of Energy and the U.S. National Science Foundation, which are gratefully thanked.

We are indebted to John Ellis and Nikolai Krasnikov for fruitful discussions.

References


[18] Parameterization of $e^-$ and $\gamma$ initiated showers in the NOMAD lead glass calorimeter, to be submitted to Nucl. Instr. and Meth.
Hadron production in diffractive deep-inelastic scattering

H1 Collaboration

M. Werner o, L.R. West c, S. Wiesand ai, T. Wilksen k, S. Willard g, M. Winde aj, G.-G. Winter k, C. Wittek m, E. Wittmann n, M. Wobisch b, H. Wollatz k, E. Wünsch k, J. Žáček af, J. Zálešák af, Z. Zhang ab, A. Zhokin y, P. Zini ad, F. Zomer ab, J. Zsembery i, M. ZurNedden al

a I. Physikalisches Institut der RWTH, Aachen, Germany 1
b III. Physikalisches Institut der RWTH, Aachen, Germany 1
c School of Physics and Space Research, University of Birmingham, Birmingham, UK 2
d Inter-University Institute for High Energies ULB-VUB, Brussels Belgium
e Universitaire Instelling Antwerpen, Wilrijk, Belgium 3
f Rutherford Appleton Laboratory, Chilton, Didcot, UK 2
g Institute for Nuclear Physics, Cracow, Poland 4
h Physics Department and IIRPA, University of California, Davis, California, USA 5
i Institut für Physik, Universität Dortmund, Dortmund, Germany 1
j DSM / DAPNIA, CEA / Saclay, Gif-sur-Yvette, France
k Department of Physics and Astronomy, University of Glasgow, Glasgow, UK 2
l DESY, Hamburg, Germany
m I. Institut für Experimentalphysik, Universität Hamburg, Hamburg, Germany 1
n II. Institut für Experimentalphysik, Universität Hamburg, Hamburg, Germany 1
o Max-Planck-Institut für Kernphysik, Heidelberg, Germany 1
p Physikalisches Institut, Universität Heidelberg, Heidelberg, Germany 1
q Institut für Hochenergiephysik, Universität Heidelberg, Heidelberg, Germany 1
r Institut für experimentelle und angewandte Physik, Universität Kiel, Kiel, Germany 1
s Institute of Experimental Physics, Slovak Academy of Sciences, Košice, Slovak Republic 6,7
t School of Physics and Chemistry, University of Lancaster, Lancaster, UK 2
u Department of Physics, University of Liverpool, Liverpool, UK 2
v Queen Mary and Westfield College, London, UK 2
w Physics Department, University of Lund, Lund, Sweden 8
x Department of Physics and Astronomy, University of Manchester, Manchester, UK 2
y CPPM, Université d’Aix-Marseille II, IN2P3-CNRS, Marseille, France
z Institute for Theoretical and Experimental Physics, Moscow, Russia
{ Max-Planck-Institut für Physik, München, Germany 1
} LAL, Université de Paris-Sud, IN2P3-CNRS, Orsay, France
[ LPNHE, École Polytechnique, IN2P3-CNRS, Palaiseau, France
\] Elinstitut für Teilchenphysik, ETH, Zürich, Switzerland 12
\][ Physik-Institut der Universität Zürich, Zürich, Switzerland 11
\][ n Institut für Physik, Humbold-Universität, Berlin, Germany 1
\][ o Rechenzentrum, Bergische Universität Gesamthochschule Wuppertal, Wuppertal, Germany 1

Received 30 March 1998
Editor: L. Montanet

Abstract

Characteristics of hadron production in diffractive deep-inelastic positron-proton scattering are studied using data collected in 1994 by the H1 experiment at HERA. The following distributions are measured in the centre-of-mass frame of
Studies of the 1992 deep-inelastic electron-proton scattering events (DIS) at HERA [1] revealed the presence of "rapidity-gap events" - events of the form $ep \rightarrow eXY$ in which the hadronic final state consists of two parts, $X$ and $Y$, separated by a large region in pseudorapidity in which no hadrons are observed. This is illustrated in Fig. 1. The masses $M_X$ and $M_Y$ of these two systems, separated by the largest rapidity gap in the event, are thus small compared to $W$, the invariant mass of the $\gamma^* p$ system. The system $Y$ in these events consists of a proton or other low-mass hadronic state and has a momentum similar to that of the incoming proton.

The magnitude of the square of the 4-momentum $p^2$ of the virtual photon interacts with a colourless space-like entity of longitudinal momentum carried by the exchange. The high-$Q^2$ virtual photon interacts with a colourless space-like entity of squared 4-momentum $t$, whose longitudinal momentum as a fraction of the target proton’s longitudinal momentum carried by the exchange. Furthermore, the pomeron may be interpreted as having partonic structure.

A QCD study of parton distribution functions [3], evolved according to the DGLAP equations [4], reveals the preference of the $F^{(3)}_2$ data for a pomeron that is dominated by a "hard-gluon" parton distribution at the starting scale of $Q^2_0 = 3\text{ GeV}^2$ (fits 2 and 3 in [3]), i.e. a distribution with a large contribution from gluons carrying a significant fraction of the momentum of the pomeron. A pomeron model with only quarks at $Q^2_0$ (fit 1 in [3]) fails.

The studies of the diffractive final state presented here may be regarded as specific tests of the pomeron structure extracted from $F^{(3)}_2$, complementary to other final-state analyses [5–7]. A gluon-dominated partonic structure of the diffractive exchange is dominated at low $Q^2$ by hard gluons.

The magnitude of the square of the 4-momentum $p^2$ of the incoming proton.

1. Introduction

Supported by the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie, FRG, under contract numbers 7AC17P, 7AC47P, 7DO55P, 7HH17I, 7HH27P, 7HD17P, 7HD27P, 7KI17I, 6MP17I, 7WT87P.

Supported by the UK Particle Physics and Astronomy Research Council, and formerly by the UK Science and Engineering Research Council.

Supported by US DOE grant DE F603 91ER40674.

Supported by the Deutsche Forschungsgemeinschaft.

Supported by the Swedish Natural Science Research Council.

Supported by Russian Foundation for Basic Researches grant no. 96-02-00019.

Supported by GA ČR grant no. 202/96/0214, GA AV ČR grant no. A1010619 and GA UK grant no. 177.

Supported by the Swiss National Science Foundation.

\footnotesize{\textsuperscript{3}Supported by the Swiss National Science Foundation.}

\footnotesize{\textsuperscript{4}Supported by the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie, FRG, under contract numbers 7AC17P, 7AC47P, 7DO55P, 7HH17I, 7HH27P, 7HD17P, 7HD27P, 7KI17I, 6MP17I and 7WT87P.}

\footnotesize{\textsuperscript{5}Supported by the UK Particle Physics and Astronomy Research Council.

\footnotesize{\textsuperscript{6}Supported by the Swedish Natural Science Research Council.

\footnotesize{\textsuperscript{7}Supported by Russian Foundation for Basic Researches grant no. 96-02-00019.

\footnotesize{\textsuperscript{8}Supported by GA ČR grant no. 202/96/0214, GA AV ČR grant no. A1010619 and GA UK grant no. 177.

\footnotesize{\textsuperscript{9}Supported by the Swiss National Science Foundation.}

\footnotesize{\textsuperscript{10}Supported by GA ČR grant no. 202/96/0214, GA AV ČR grant no. A1010619 and GA UK grant no. 177.

\footnotesize{\textsuperscript{11}Supported by the Swiss National Science Foundation.}
pomeron is expected to interact largely by boson-gluon fusion (BGF); the transverse momenta of the outgoing partons relative to the photon direction in the $\gamma^* p$ centre-of-mass (CM) frame (i.e. the CM frame of the photon dissociation system $X$), as well as the amount of gluon radiation and consequently the parton multiplicity, are therefore likely to be greater than in the case of a quark-dominated pomeron.

The partonic structure of the pomeron should thus be reflected in the energy-flow and charged-particle distributions in the $\gamma^* p$ CM frame. These distributions are compared here with data on deep-inelastic $\mu p$ interactions in the $\gamma^* p$ CM frame at the EMC experiment. This provides a comparison of the structure of the pomeron with that of the proton.

The distributions are also compared with predictions from a calculation using a pomeron with parton distributions from the best QCD fit to $F^D_2(x)$ (fit 3). Comparison with predictions from a calculation with only quarks in the pomeron at the starting scale of the QCD fit (fit 1) demonstrates the sensitivity of the measurements to the parton distributions of the pomeron.

Another approach to diffractive DIS is provided by the photon dissociation picture [8]. In the rest frame of the proton, the photon fluctuates a long time before the interaction into a Fock state with definite parton content ($qg, q\bar{q}, \ldots$), and this partonic state scatters diffractively off the proton. Scattering from a quark in the pomeron corresponds to an Aligned Jet Model [9] topology in this picture, whereas scattering from a gluon corresponds to Fock states with one or more gluons and results in less pronounced alignment.

A non-diffractive model for the production of events with a large rapidity gap has also been proposed, using a conventional picture of DIS based on scattering from a single parton within the proton, followed by soft colour interactions [10,11]. The data are also compared with predictions of this model.

2. Detector, event selection and kinematics

The data used here were collected in 1994 with the H1 detector at the HERA collider, which operated with 27.5 GeV positron and 820 GeV proton beams. The H1 detector is described in detail elsewhere [12]; those components of importance for the analyses presented here are briefly mentioned in the following. The laboratory coordinate system has its origin at the nominal interaction point and its $z$ axis in the proton beam direction, also called the forward direction. The pseudorapidity is defined as $\eta = -\ln \tan \frac{\theta}{2}$, where $\theta$ is the angle with respect to the $z$ axis.

In the forward and central regions the interaction point is surrounded by a system of tracking detectors – interleaved drift and multi-wire proportional chambers – which cover the pseudorapidity range $-1.5 < \eta < 2.8$ and the full azimuth. The momenta of charged particles are determined for this analysis from their track curvature in the central jet chamber (CJC) in the uniform magnetic field of strength 1.15 T generated by a superconducting solenoid. The CJC has a resolution of $\sigma_p/p_T \approx 0.006 p_T \oplus 0.008$ (where $p_T$ is in GeV and the constant term describes the contribution from multiple scattering at $p_T = 0.5 \text{ GeV}$ and $\sigma_\theta = 20 \text{ mrad}$). The solenoid surrounds the trackers and the fine-grained liquid-argon (LAr) calorimeter, which covers the range $-1.5 < \eta < 3.6$ and is used to measure energies in the hadronic final state with a resolution of $\sigma_{E}/E \approx 0.5/\sqrt{E} \oplus 0.02$ (where $E$ is in GeV) [12]. Charged-particle detection in the backward region, $-3.1 < \eta < -1.5$, is provided by the backward proportional chamber (BPC). Behind this, the backward electromagnetic calorimeter (BEMC) completes the calorimetric coverage in the range $-3.5 < \eta < -1.5$.

Several subdetectors in the forward region are used in this analysis to tag particles emitted close to the proton direction. These detectors cover a range larger than their purely geometrical acceptance, due to the effect of secondaries resulting from the scattering of primary particles in the beam pipe and adjacent material. The copper-silicon plug calorimeter covers the range $3.5 < \eta < 5.5$. The first three double layers of drift chambers in the forward muon detector (FMD) are sensitive to particles in the range $5.0 < \eta < 6.5$. The proton-remnant tagger (PRT), comprising a set of double layers of scintillators situated around the proton beam pipe, covers the range $6.0 < \eta < 7.5$.

The luminosity is determined from the rate of Bethe-Heitler interactions detected in a monitor
There must also be no energy deposit of more than \( w_x \) in the forward components of the H1 detector. This gap is tagged by the absence of signals above \( Y \). These variables are defined as

\[
Q^2 = -q^2, \quad y = (q \cdot p)/(k \cdot p) = (W^2 + Q^2)/s, \quad \text{where } p, k \quad \text{and} \quad q \quad \text{are the 4-momenta of the incoming proton, positron and photon respectively, } W \quad \text{is the CM energy of the } \gamma^* p \quad \text{system, and } s \quad \text{is the squared CM energy of the } ep \quad \text{system. The variables } Q^2 \quad \text{and } y \quad \text{are calculated from the measured energy and angle of the scattered positron.}
\]

Diffractive events are recognised by the rapidity gap between the outgoing hadronic systems \( X \) and \( Y \). This gap is tagged by the absence of signals above noise in the forward components of the H1 detector [3,14]: the plug calorimeter, the FMD and the PRT. There must also be no energy deposit of more than 400 MeV in the forward region (\( \eta > 3.0 \)) of the LAr calorimeter.

The variable \( x_p \) is defined as

\[
x_p = \frac{q \cdot P}{q \cdot p}
\]

where \( P \) is the 4-momentum exchanged between systems \( X \) and \( Y \), and is reconstructed using the relation

\[
x_p = \frac{1}{2E_p} \sum_{e+X} (E + p_e).
\]

Here, \( E_p \) is the energy of the incoming proton, \( E \) and \( p_e \) are the energy and longitudinal momentum of each final-state particle in the laboratory frame, and the sum runs over the scattered positron \( e \) and all detected particles in the photon dissociation system \( X \). The particles are reconstructed using a combination of tracks and calorimeter clusters, with an algorithm for track-cluster association avoiding double counting [15]. In order to enhance the contribution from pomeron exchange and to reduce the contribution from meson exchange, the requirement \( x_p < 0.025 \) is made. The fraction of the exchanged momentum \( P \) carried by the struck quark is given by \( \beta \), where

\[
\beta = \frac{Q^2}{2q \cdot P} = \frac{Q^2}{Q^2 + M_X^2}.
\]

Neither the squared momentum transfer \( t \) nor the mass \( M_X \) of the hadronic system \( Y \) is measured here. However, the requirement of the absence of activity in the forward detectors imposes the approximate restrictions \( M_X < 1.6 \text{GeV} \) and \( |t| < 1 \text{GeV}^2 \), and the results are corrected to this kinematic region. Since \( M_X \) is not measured, it is not possible to distinguish events containing an elastically scattered proton from those in which the proton dissociates into a low-mass state.

The analyses are performed in the \( \gamma^* P \) CM frame, the 4-momentum of the \( \gamma^* P \) system being reconstructed as \( V = q + x_p p \) [16]. The transverse momentum of the pomeron with respect to the direction of the incoming proton is not measured, but can be neglected since this has no significant effect on the measurements presented here.

The energy-flow distribution is measured using energy deposits reconstructed from clusters of cells in the LAr and BEMC. The clusters are required to have an energy in the range \( 0 < E < 25 \text{GeV} \) and to lie in the angular range \(-3 < \eta < 4\). The charged-particle distributions are measured using only particles that are detected in the CJC, originate from the \( ep \) interaction vertex in the angular region \(-1.31 < \eta < 1.31\), and have a transverse momentum in the laboratory frame of \( 0.15 < p_T < 10 \text{GeV} \), but the results are fully corrected with Monte Carlo simulations (see below) to cover the whole hadronic system.
Two variables are calculated for each charged particle, both defined in the CM frame of the γ p system. The Feynman-x variable, \( x_F \), is defined as

\[
2 p^+_{\gamma} \, x_F = \frac{2 p^+_{\gamma}}{M_X},
\]

where \( p^+_{\gamma} \) is the component of the particle’s momentum in the direction of motion of the incoming photon. In non-diffractive DIS, \( x_F \) is defined in the \( \gamma p \) CM frame, and the denominator is \( W \). Positive \( x_F \) corresponds to the current hemisphere and negative \( x_F \) to the target fragmentation hemisphere. The other variable used here, \( p_T \), is the transverse momentum of the particle with respect to the photon direction.

3. Monte Carlo models for diffractive interactions

Monte Carlo calculations employing the event generator RAPGAP 2.02 [17] are used in conjunction with a detailed simulation of the H1 apparatus to correct the experimental distributions for the acceptance and resolution of the detector, and later to compare the results with theoretical expectations. The data are also compared with the predictions of the Monte Carlo generator LEPTO 6.5 [11].

The RAPGAP model treats diffractive interactions as inelastic \( eP \) collisions, the pomeron being modelled as an object with partonic substructure. Scattering on mesons is also included.

For scattering from quarks, the lowest-order diagram considered by RAPGAP at the parton level is \( \sigma(q \bar{q}) \) scattering (Fig. 2a), a higher-order, \( \sigma(\alpha \alpha) \) process is QCD Compton (QCD-C) scattering \( eq \rightarrow eq \) (Fig. 2b). For scattering from gluons, the lowest-order \( \sigma(\alpha \alpha) \) process is boson-gluon fusion \( eg \rightarrow eqg \) (Fig. 2c). A cut \( \hat{p}_T^2 > 2 \text{GeV}^2 \) is applied in order to avoid divergences in the \( \sigma(\alpha \alpha) \) matrix elements for massless quarks, \( \hat{p}_T \) being the transverse momentum of the outgoing partons with respect to the photon direction in the CM frame of the hard subprocess. For \( \hat{p}_T^2 < 2 \text{GeV}^2 \), only \( eq \rightarrow eq \) scattering is used. In the RAPGAP model, a “pomeron remnant” is implicit at the parton level, consisting of a quark (antiquark) in \( eq \) (\( eqg \)) scattering and a gluon in the \( eg \) case (see Fig. 2). The renormalisation and factorisation scale \( \mu^2 \) is set to \( Q^2 \). Higher-order corrections at the parton level are treated using leading-log parton showers (PS) [18] as implemented in RAPGAP. Subsequent hadronisation is simulated according to the Lund string model in JETSET [19]. The dependence of the acceptance corrections on the method used to handle higher-order corrections in RAPGAP is studied with a separate calculation in which QCD-C and higher-order processes are simulated by the colour-dipole (CD) model.

![Fig. 2. Elementary processes included in the RAPGAP simulation of \( eP \) diffractive interactions: (a) lowest-order, \( \sigma(q) \), \( e \) scattering from a quark, leading to a two-parton configuration; (b) one of the diagrams for the QCD-C process, \( \sigma(\alpha \alpha) \); (c) lowest-order, \( \sigma(\alpha \alpha) \), \( e \) scattering from a gluon by BGF, leading to a three-parton configuration.](image-url)
RAPGAP is interfaced to HERACLES [21] for the simulation of QED radiative effects.

For the comparisons presented here, two sets of parton distributions for the pomeron are used in the RAPGAP generator. The first is taken from the best fit to the \( Q^2 \) and \( \beta \) dependence of the diffractive structure function \( F_2^{D}(x) \) (fit 3 in [3]). It is characterised by a hard gluon distribution at the starting scale for DGLAP evolution, \( Q_0^2 = 3 \text{ GeV}^2 \), in which gluons carry \( \geq 80\% \) of the total momentum of the diffractive exchange. The second is taken from a fit in which only quarks are permitted to contribute to the partonic structure of the pomeron at \( Q^2 = Q_0^2 \) (fit 1 in [3]). The latter choice does not provide a satisfactory description of the \( F_2^{D}(x) \) measurements; it is used here to demonstrate the sensitivity of the results to the parton distributions of the pomeron. The predictions of RAPGAP with these two sets of parton distributions are hereafter referred to as RG±F fit 3 and RG±F fit 1 respectively. The partons are treated as having no intrinsic \( k_T \).

In the LEPTO 6.5 model, rapidity-gap events are generated by the soft colour interaction (SCI) mechanism. Deep-inelastic scattering is modelled using the matrix elements for processes up to \( \sigma(\alpha\alpha_s) \), as in RAPGAP, but with partons coming directly from the proton according to the MRS(H) parameterisation [23] of the proton structure function, in which the renormalisation and factorisation scale \( \mu^2 \) is again set to \( Q^2 \). The divergences in the \( \sigma(\alpha\alpha_s) \) matrix elements are avoided using the cuts \( \delta > \delta_{\text{min}} \) and \( z_q < z_{q,\text{min}} < 1 - z_{q} \), where \( \delta \) is the CM energy of the hard subprocess, \( z_q = (p \cdot p_q)/(p \cdot q) \) and \( p_q \) is the 4-momentum of one of the outgoing partons. The parameters \( \delta_{\text{min}} \) and \( z_{q,\text{min}} \) are set to 4 GeV\(^2\) and 0.04 respectively. Higher-order corrections are treated with the PS method, and hadronisation follows the Lund string model. Further non-perturbative interactions take place as the outgoing partons pass through the colour field of the proton. These soft colour interactions can result in a hadronic final state comprising two colour-singlet systems separated by a rapidity gap.

4. Acceptance corrections and systematic errors

The data are corrected for the acceptance and resolution of the H1 apparatus using events generated by RAPGAP 2.02 [17] with a hard-gluon pomeron structure function. The following sources of systematic error are taken into account; the errors shown in brackets are those on the energy flow and the charged particle distributions respectively:

- an uncertainty in the LAr calibration for hadronic energy of \( \pm 4\% \), leading to an average error in the experimental distributions of \( \pm (6\%, 2\%) \);
- an uncertainty in the BEMC calibration for hadronic energy of \( \pm 20\% \), leading to an average error of \( \pm (6\%, 0.7\%) \);
- an uncertainty in the \( 1\% \) in the positron energy scale, leading to an average error of \( \pm (3\%, 0.5\%) \);
- an uncertainty of \( \pm 1 \text{ mrad} \) in the positron polar angle, leading to an average error of \( \pm (0.5\%, 0.5\%) \);
- an uncertainty in the \( x_p \) dependence in the Monte Carlo used to calculate the acceptance corrections, taken into account by reweighting the generated events with the function \( x_p^{0.2} \), resulting in an average error of \( \pm (8\%, 3\%) \);
- an uncertainty in the \( \beta \) dependence in the Monte Carlo, taken into account by reweighting the generated events with the function \( (a^{-1} - a)\beta + a \), where \( a \) has the range 0.5 to 2, resulting in an average error of \( \pm (6\%, 1\%) \);
- an uncertainty in the \( t \) dependence in the Monte Carlo, taken into account by reweighting the generated events with the function \( e^{\pm 2t} \), resulting in an average error of \( \pm (5\%, 2\%) \);
- uncertainties arising from the method used to treat higher-order corrections in the Monte Carlo; this results in an error of \( \pm (10\%, 10\%) \), evaluated by taking the full difference between the results obtained using the PS and CD models as an estimation of the extrapolation uncertainty from the Monte Carlo.

\(^{22}\) Fits 2 and 3 in [3] give very similar predictions for the distributions studied here. Only the predictions from fit 3 are shown in this paper.
In the charged-particle analysis, the track selection criteria were varied to allow for the imperfect description of the Central Tracker in the Monte Carlo. The selection was varied by increasing the angular range of the tracks to $-1.50 < \eta < 1.64$ and by requiring the radial track length to be greater than 15 cm and the number of hits on a track $N_{\text{hits}}$ to be greater than 10. This results in an average error of $\pm 5\%$, rising in the low-$x_F$ region to $20\%$ in the $x_F$ distribution and $15\%$ in $\langle p_T^2 \rangle$ in the ‘seagull plot’ (Fig. 6).

5. Results

Results $^{13}$ are shown for the kinematic range $7.5 < Q^2 < 100 \text{GeV}^2$, $0.05 < y < 0.6$, $x_p < 0.025$, $|t| < 1 \text{GeV}^2$ and $M_X < 1.6 \text{GeV}$. All distributions are fully corrected for the effects of the acceptance and resolution of the H1 apparatus, and statistical and systematic errors on the data points are combined quadratically. The inner vertical error bars indicate the statistical error and the outer bars the total error. Due to the $x_p$ cut, the meson-exchange contribution is small ($< 7\%$ in the RG-$F_2^D$ (fit 3) model) and does not affect the conclusions.

It is shown below that, in Monte Carlo calculations, the characteristic features of the parton content of the pomeron are reflected in the distributions of various final-state observables. The data are therefore compared with data on $\gamma^*P$ interactions at the EMC experiment $^{24}$, with the mean $W$ of the EMC inclusive DIS data similar to the mean $\gamma^*P$ CM energy, $\langle M_X \rangle \approx 12 \text{GeV}$, of the H1 diffractive data.

The experimental data are also compared with the predictions of the RAPGAP model with the two different pomeron structures from the DGLAP fits to $F_2^{\text{QCD}}$ – the preferred hard-gluon structure (RG-$F_2^D$ (fit 3)) and the structure with only quarks at the starting scale (RG-$F_2^Q$ (fit 1)) – as well as with the soft colour interaction (SCI) model as implemented in LEPTO 6.5.

In the CM frame of the $\gamma^*P$ system, a model with a quark-dominated pomeron results in events in which the struck quark and the pomeron remnant tend to be close to the $\gamma^*P$ axis. Large contributions to the transverse momentum can arise from QCD corrections such as QCD Compton scattering, but the rate of these is suppressed by a factor $\alpha_s$. In contrast, if the pomeron is a gluon-dominated object, more transverse momentum and energy flow are produced at lowest order. This arises because the quark propagator in the BGF process can have non-zero virtuality, so the quark-antiquark pair from the hard subprocess is not necessarily aligned along the $\gamma^*P$ axis. Evidence for this effect is seen in a recent analysis, using the same data, of the distribution of the $p_T$ of thrust jets relative to the $\gamma^*P$ axis $^{[6]}$. In both the quark- and gluon-dominated cases, small contributions to the transverse momentum are generated by the intrinsic $k_T$ of the partons, by the hadronisation process, and by particle decays. However, it is known from the average thrust distribution shown in $^{[6]}$ that a quark-dominated exchange, even with large intrinsic $k_T$, would not be able to explain the relatively low value of the average thrust.

Fig. 3 shows the event-normalised energy flow $1/N \int dE/d\eta^*$ in three different regions of $M_X$, $\eta^*$ being the pseudorapidity relative to the direction of motion of the incoming photon in the CM frame of the $\gamma^*P$ system and $N$ being the number of events. The distribution is approximately symmetrical about $\eta^* = 0$, with similar levels of energy flow in the two hemispheres. At higher masses, $M_X > 8 \text{GeV}$, a two-peak structure is seen, indicating that the major topological property of these events is of a 2-jet nature (current jet and remnant jet).

The data are well described by the RG-$F_2^D$ (fit 3) model $^{14}$, and for $M_X > 8 \text{GeV}$ by LEPTO 6.5, whereas the RG-$F_2^Q$ (fit 1) model predicts too much energy flow at the largest accessible pseudorapidity.

$^{13}$ The data are available in numerical form on request and have also been submitted to the Durham HEPDATA database http://durpdg.dur.ac.uk/HEPDATA.

$^{14}$ The $\beta_T$ cut in the generator was varied from 2 up to 4 GeV$^2$ without affecting the conclusions drawn from the distributions presented in this paper.
and fails to account for the observed energy flow in the central region, $\eta^+ = 0$.

These features can be related to the gluon content of the pomeron and to the role of BGF. The transverse momentum generated by this process reduces the accessible pseudorapidity range along the $\gamma^+P$ axis at fixed $M_X$ and induces additional energy flow in more central regions. In addition, more energy is expected in the central region for a gluon-dominated than for a quark-dominated object because of the exchange of a gluon in BGF; this results in a colour octet-octet field between the outgoing quarks and the gluonic remnant (which all together form an overall colour singlet), giving rise to enhanced soft gluon radiation compared to the triplet-antitriplet field between the struck quark and the remnant in $e\gamma$ scattering [16,25]. This effect has been seen for gluon jets produced in $e^+e^-$ interactions [26]. This enhancement is also apparent in the virtual photon dissociation picture [8], where the process corresponding to BGF involves a Fock state containing a low-energy gluon.

The production of charged particles is studied using the distributions of the variables $x_F$ and $p_T^2$, defined in Section 2. The results are shown for the restricted mass range $8 < M_X < 18 \text{ GeV}$ in order to...
make the data more directly comparable with the EMC $\gamma^* p$ results.

The $x_F$ distribution, normalised by the number of events $N$, is shown in Fig. 4. A feature of the diffractive data is the similarity between the $x_F$ distributions in the two hemispheres, in contrast to the strong asymmetry of the $\gamma^* p$ data. The asymmetry in the latter case is explained by the requirement of baryon number conservation in $ep$ interactions. This results in a large contribution from baryons in the region $x_F \leq -0.4$ [24].

The $x_F$ distribution is well described by the RG--$F_1^D$ (fit 3) and LEPTO 6.5 models, whereas the RG--$F_2^D$ (fit 1) calculation fails in that it predicts too little particle production in the central region ($x_F \approx 0$) and too much particle production at large $|x_F|$. These features are related to those observed in the energy-flow distribution: longitudinal phase space is restricted because of the transverse momentum generated by the BGF process, and there is additional radiation in the central region for a gluon-dominated object.

Fig. 4. The Feynman-$x$ ($x_F$) distribution, showing H1 $\gamma^* p$ data in the $\gamma^* p$ CM frame in the kinematic region $7.5 < Q^2 < 100 \text{GeV}^2$, $0.05 < y < 0.6$, $x_F < 0.025$, $\Delta x < 18 \text{GeV}$, $|t| < 1 \text{GeV}^2$ and $M_t < 1.6 \text{GeV}$, together with EMC $\mu p$ DIS data in the $\gamma^* p$ CM frame. Positive $x_F$ corresponds to the direction of motion of the incoming photon. Each point is plotted at the horizontal position where the predicted probability density function has its mean value in the bin, as prescribed in [27]. The uncertainty in the horizontal position due to the model-dependence of the shape of the distribution is indicated by horizontal bars, but in most cases is contained within the width of the symbol. The bin boundaries are marked by vertical dotted lines. Also shown are the predictions of the following Monte Carlo models for the H1 data: RAPGAP with the hard-gluon pomeron structure taken from fit 3 in [3] (RG--$F_1^D$ (fit 3)); RAPGAP with the pomeron structure containing only quarks at the starting scale, taken from fit 1 in [3] (RG--$F_2^D$ (fit 1)); and the soft colour interaction model as implemented in LEPTO 6.5.
The $p_T^2$ distribution, normalised by the number of events $N$, is shown in Fig. 5. Results are shown for the range $0.2 < x_F < 0.4$. This restricts the data to the current region, where a comparison between H1 and EMC data is more meaningful, and matches the $x_F$ range chosen by EMC. The number of high-$p_T$ charged particles ($p_T^2 \geq 2.0$ GeV$^2$) is significantly higher in the $\gamma^* P$ data than in the EMC $\gamma^* p$ data. This points to a larger contribution from scattering from gluons in the diffractive case than in inclusive DIS.

The $p_T^2$ distribution is well described by the RG-$F^D$ (fit 3) calculation, which predicts a harder spectrum than the RG-$F^D$ (fit 1) model because of the role of the BGF process. The data are also well described by LEPTO 6.5. The RG-$F^D$ (fit 1) model does not describe the distribution well at high $p_T^2$. It should also be noted that the RG-$F^D$ (fit 1) predictions for the $p_T^2$ distribution are in good agreement with the EMC $\gamma^* p$ data. This supports the argument that the EMC data (at a mean $Q^2$ of 12 GeV$^2$) are dominated by quarks and that the

![Fig. 5. The $p_T^2$ distribution, showing H1 $\gamma^* P$ data in the $\gamma^* P$ CM frame in the kinematic region $7.5 < Q^2 < 100$ GeV$^2$, $0.05 < y < 0.6$, $x_F < 0.025$, $8 < M_T < 18$ GeV, $|t| < 1$ GeV$^2$, $M_E < 1.6$ GeV, and $0.2 < x_F < 0.4$, together with EMC $\mu p$ DIS data in the $\gamma^* p$ CM frame. Each point is plotted at the horizontal position where the predicted probability density function has its mean value in the bin, as prescribed in [27]. The uncertainty in the horizontal position due to the model-dependence of the shape of the distribution is indicated by horizontal bars, but in most cases is contained within the width of the symbol. The bin boundaries are marked by vertical dotted lines. Also shown are the predictions of the following Monte Carlo models for the H1 data: RAPGAP with the hard-gluon pomeron structure taken from fit 3 in [3] (RG-$F^D$ (fit 3)); RAPGAP with the pomeron structure containing only quarks at the starting scale, taken from fit 1 in [3] (RG-$F^D$ (fit 1)); and the soft colour interaction model as implemented in LEPTO 6.5.](image-url)
higher $p_T^2$ in the diffractive data is not an effect of the larger $Q^2$ at H1 (where the mean $Q^2$ in this analysis is 25 GeV$^2$).

Confirmation of the general trends discussed above is obtained from the "seagull plot", shown in Fig. 6, in which the mean transverse momentum squared, $\langle p_T^2 \rangle$, is plotted as a function of $x_F$. It is observed that $\langle p_T^2 \rangle$ is significantly higher in the H1 diffractive data than in the EMC $\gamma^* p$ data at all but the highest $x_F$ values, confirming the conclusions drawn from the comparisons in Fig. 5. The seagull plot also shows a greater degree of symmetry than is present in the $\gamma^* p$ data (cf. Fig. 4). As well as the effect of baryon number conservation, this is indicative of more parton radiation in the target fragmentation region. This is consistent with having a more point-like partonic system in the target fragmentation hemisphere in the diffractive data than in the $\gamma^* p$ data, where the extended nature of the proton remnant leads to a restricted phase space for parton radiation. Note, however, that there might be an indication of a small asymmetry between the two hemispheres in the energy-flow distribution (Fig. 3).

The significant increase in $\langle p_T^2 \rangle$ as $|x_F|$ increases from 0 to $\sim 0.5$ is well described by the RG–Fi$^2$ (fit 3) calculation and reasonably well de-

![Fig. 6. Seagull plot showing H1 $\gamma^* p$ data in the $\gamma^* p$ CM frame in the kinematic region $7.5 < Q^2 < 100\text{GeV}^2$, $0.05 < y < 0.6$, $x_F < 0.025$, $|t| < 1\text{GeV}^2$, $M_p < 1.6\text{GeV}$ and $8 < M_\gamma < 18\text{GeV}$, together with EMC $\mu p$ DIS data in the $\gamma^* p$ CM frame. Positive $x_F$ corresponds to the direction of motion of the incoming photon. The points are plotted at the centre of the bin in the horizontal coordinate, and the horizontal bars indicate the width of the bin. Also shown are the predictions of the following Monte Carlo models for the H1 data: RAPGAP with the hard-gluon pomeron structure taken from fit 3 in [3] (RG–Fi$^2$ (fit 3)); RAPGAP with the pomeron structure containing only quarks at the starting scale, taken from fit 1 in [3] (RG–Fi$^0$ (fit 1)); and the soft colour interaction model as implemented in LEPTO 6.5.](image-url)
scribed by LEPTO 6.5, whereas the \( \left< p_T^2 \right> \) values predicted by the RG–F2P (fit 1) model are too low across most of the \( x_F \) range, in accordance with the above discussion of Fig. 5.

6. Summary and conclusions

Energy-flow and charged-particle spectra have been measured in diffractive deep-inelastic scattering at HERA in the centre-of-mass (CM) frame of the photon dissociation (\( \gamma^* \gamma \)) system. The data support the conclusion reached in the analysis of the diffractive structure function \( F_2^{D(3)} \) that, at low \( Q^2 \), the momentum of the diffractive exchange is carried largely by hard gluons. Thus, significant transverse momentum and energy flow are produced by the hard subprocess in the boson-gluon fusion mechanism, and also arise from the enhanced soft gluon radiation associated with the gluon from the diffractive exchange. In the photon dissociation picture, the same data can correspondingly be interpreted as evidence for a significant contribution from events with photon fluctuation into Fock states with one or more gluons.

The data have been compared with the \( \gamma^* p \) data of the EMC collaboration, with the \( \gamma^* p \) CM energy (\( W \)) of the EMC inclusive DIS data similar to the \( \gamma^* \gamma \) CM energy (\( M_X \)) of the H1 diffractive data. It is seen that additional transverse momentum is produced in diffractive scattering compared to the \( \gamma^* p \) case. A striking contrast is observed in the seagull plot, indicating an approximate symmetry between the target and current hemispheres in diffractive DIS, as would be expected in the hard-gluon picture of the pomeron and in the photon dissociation picture, whereas radiation is significantly suppressed in the proton-remnant region in non-diffractive DIS.

The features of hadron production are well reproduced by a model featuring a factorisable pomeron flux in the proton if the partonic structure of the pomeron is dominated by hard gluons at the starting scale of \( Q^2_0 = 3 \text{ GeV}^2 \) used in a DGLAP analysis of the diffractive structure function \( F_2^{D(3)} \). In contrast, a model with a pomeron consisting, at low \( Q^2 \), primarily of quarks fails to describe the data in several respects. With respect to the \( \gamma^* \gamma \) axis, the quark-based pomeron produces too little energy flow and particle production in the central region, but too much at large \( |\eta| \) and \( |x_F| \). In the \( p_T^2 \) and “seagull” distributions, this model produces too little transverse momentum. A model based on soft colour interactions (SCI) also gives an acceptable description of the data.

Acknowledgements

We are grateful to the HERA machine group, whose outstanding efforts have made and continue to make this experiment possible. We thank the engineers and technicians for their work in constructing and maintaining the H1 detector, our funding agencies for financial support, the DESY technical staff for their continual assistance, and the DESY directorate for the hospitality which they extend to the non-DESY members of the collaboration.

References

Nucleon-deuteron scattering from an effective field theory

P.F. Bedaque a,1, U. van Kolck b,2,3

a Institute for Nuclear Theory, University of Washington, Seattle, WA 98195-1560, USA
b Department of Physics, University of Washington, Seattle, WA 98195-1560, USA

Received 3 November 1997; revised 6 March 1998
Editor: J.-P. Blaizot

Abstract

We use an effective field theory to compute low-energy nucleon-deuteron scattering. We obtain the quartet scattering length using low energy constants entirely determined from low-energy nucleon-nucleon scattering. We find $a_{s} = 6.33$ fm, to be compared to $a_{s} = 6.35 \pm 0.02$ fm.

There has been considerable interest lately in a description of nuclear forces from the low-energy effective field theory (EFT) of QCD. (For a review, see Ref. [1].) Following a program suggested by Weinberg [2], the leading components of the nuclear potential have been derived [3] and a reasonable fit to two-nucleon properties has been achieved [4]. The correct formulation of the nuclear force problem within the EFT method is important because it will allow a systematic calculation of nuclear properties consistently with QCD. One would like, for example, to devise a theory of nuclear matter rooted in a hadronic theory that treats chiral symmetry correctly and yields the well-known few-nucleon phenomenology. One hopes that after a number of parameters of the EFT are either calculated from first principles or fitted to a set of few-nucleon data, the theory can be used to predict other reactions involving light nuclei and features of heavier nuclei.

However, some issues concerning renormalization in this non-perturbative context and fine-tuning in the two-nucleon $S$-waves have been raised in Refs. [5,6] and are still not fully resolved [7,8]. The fine-tuning necessary to bring a (real or virtual) bound state very close to threshold generates a scattering length $a$ much larger than other scales in the problem. At momenta of $O(1/a)$, mesons can be integrated out and the characteristic mass scale $\mu$ of the underlying theory controls the size of the other effective range parameters; for example, the effective range $r_{0} \sim 2/\mu$. Once the leading order contributions, which give rise to $a$, are included to all orders, the EFT at momenta $O(1/a)$ becomes an expansion in powers of $1/(a\mu)$. Kaplan [6] has noticed that the interactions that generate a non-zero $r_{0}$ can also be resummed by the introduction of a baryon number

---

1 E-mail: bedaque@phys.washington.edu.
2 E-mail: vankolck@phys.washington.edu
3 Address after Jan 1 1998: Kellogg Radiation Lab 106-38, California Institute of Technology, Pasadena, CA 91125; E-mail: vankolck@krl.caltech.edu.
two state of mass $\Delta = 2/M\rho_0$, which in lowest order in a derivative expansion couples to two nucleons with a strength $g^2/4\pi = 1/M^2|\rho_0|$. In Ref. [6] it was shown how this works in the two-nucleon $^3S_1$ channel. Analogous considerations hold for the $^3S_1$ channel, where they are similar to the old quasi-particle approach of Weinberg [9].

In this paper we consider the application of these ideas to the three-nucleon system. Our goal here is to calculate some of the three-nucleon parameters that are dominated by the leading interactions in the EFT without pions. We show in particular that the quartet scattering length in neutron-deuteron scattering can be predicted once the EFT is constrained by low-energy two-nucleon data. Such an attempt to a model-independent or ‘universal’ approach is not a new idea; it permeates for example the work of Efimov see, e.g., [10] and Amado see, e.g., [11]. However, as we will show, the EFT formulation is much easier to implement, from both conceptual and practical standpoints.

For momenta of order $1/a$ (the momentum scale relevant for zero-energy $Nd$ scattering), we can integrate out mesons and consider an EFT with only nucleons $N$. Interactions are then described by a tower of nucleon contact operators with an increasing number of derivatives. Amplitudes in leading order are given by a zero-range four-nucleon interaction iterated to all orders. Corrections come in powers of $1/a^2$. The next two orders in this expansion, $1/(am_\pi) \sim r_0/2a$ and $1/(am_\pi)^2 \sim (r_0/2a)^2$, stem from one and two insertions of a two-derivative four-nucleon operator giving rise to a non-zero $r_0$. It is advantageous to sum all the contributions coming from this operator, which can be easily done because they appear in a geometric series. The resulting interaction is equivalent to the $s$-channel propagation of a dibaryon, and therefore can be obtained more directly by the introduction of a dibaryon field.

Since in both $I = 0$ and $I = 1$ $S$-wave two-nucleon channels we observe one real, one virtual bound state near threshold, we consider two dibaryon fields, $T$ ($D$) of spin zero (one) and isospin one (zero). The most general Lagrangian invariant under parity, time-reversal, and small Lorentz boosts is

$$\mathcal{L} = N^\dagger \left( i\partial_0 + \frac{\mathbf{\nabla}^2}{2M} + \ldots \right) N + T^\dagger \cdot \left( -i\partial_0 - \frac{\mathbf{\nabla}^2}{4M} + \Delta_T + \ldots \right) T + D^\dagger \cdot \left( -i\partial_0 - \frac{\mathbf{\nabla}^2}{4M} + \Delta_D + \ldots \right) D$$

$$- \frac{g_T}{2} \left( T^\dagger \cdot N\sigma_2 \tau_2 \tau_2 N + \text{h.c.} \right) - \frac{g_D}{2} \left( D^\dagger \cdot N\sigma_2 \sigma_2 \tau_2 N + \text{h.c.} \right) + \ldots$$

(1)

Here the $\Delta_T, \Delta_D$ and $g_T, g_D$ are undetermined parameters and ‘‘…’’ stands for higher order terms. (Note that the effects of non-derivative and two-derivative four-nucleon terms can be absorbed into a redefinition of $\Delta_T, \Delta_D$ and $g_T, g_D$ and higher order four-nucleon terms.)

In this non-relativistic theory all particles propagate forward in time, nucleon tadpoles vanish and, as a consequence, there is no dressing of the nucleon propagator, which is simply

$$S_N(p) = \frac{i}{p^0 - \frac{p^2}{2M} + i\epsilon}.$$  

(2)

The propagators for dibaryons are more complicated, because of the coupling to two-nucleon states. The dressed propagators consist of the bubble sum in Fig. 1, which amounts to a self-energy contribution proportional to the

\[ \text{Fig. 1. Dressed dibaryon propagator.} \]
bubble integral. This integral is proportional to the (large) mass \( M \), and it is this enhancement that gives rise to non-perturbative phenomena and leads eventually to the existence of bound states. The integral is also ultraviolet divergent and requires regularization. Introducing a cut-off \( \Lambda \) we find a linear divergence \( \propto \Lambda \), a cut-off independent piece which is non-analytic in the energy, a term that goes as \( \Lambda^{-1} \) and terms that are higher order in \( \Lambda^{-1} \). The first and third terms can be absorbed in renormalization of the parameters of the Lagrangian \( \mathcal{L} \); in what follows we omit a label \( R \) that should be attached to these parameters, i.e., \( \Delta_{T,D} \) and \( g_{T,D} \) stand for renormalized parameters. Higher order terms are neglected because they are of the same order as interactions in the “…” of the Lagrangian \( \mathcal{L} \). A dibaryon propagator has therefore the form

\[
iS_D(p) = \frac{1}{p^0 - p^2/4M - \Delta_D + \left( M g_D^2 / 2\pi \right) \sqrt{-Mp^0 + p^2/4 - \epsilon + i\epsilon}}.\tag{3}
\]

Note that such a dressed propagator has two poles at \( p^0 = p^2/4M - B \) and \( p^0 = p^2/4M - B_{\text{deep}} \) and a cut along the positive real axis starting at \( p^0 = p^2/4M \).

The NN amplitude can now be obtained directly from \( S_R(p) \) as in Fig. 2. In the center-of-mass, the on-shell \( I = 0 \), \( J = 1 \) S-wave amplitude at an energy \( E = k^2/M \) is

\[
3T_{NN}(k) = \frac{4\pi}{M} \frac{1}{\left( M g_D^2 / 2\pi \right) + \left( M^2 S_D^2 / 2\pi \right) k^2 - ik},\tag{4}
\]

which is exactly equivalent to the effective range expansion. An analogous result holds for the \( I = 1 \) S-wave. The four parameters \( \Delta_{T,D} \) and \( g_{T,D} \) can then be fixed from the experimentally known scattering lengths and effective ranges. The NN amplitude has shallow poles at \( B \approx 1/\sqrt{M} \) which are associated with the deuteron in the \( ^3S_1 \) channel and with the virtual bound state in the \(^1S_0 \) channel. The effective theory has also an additional deep bound state in each channel at \( B \approx 4/\sqrt{M} \), which is outside the range of validity of the EFT.

From the triplet parameters \( a_s = 5.42 \) fm and \( r_s = 1.75 \) fm [12] we find \( \Delta_D = 8.7 \) MeV and \( g_D^2 = 1.6 \cdot 10^{-3} \) MeV\(^{-1}\). The resulting deuteron binding energy is \( B = 2.28 \) MeV. From the singlet parameters \( a_{np} = -17.3 \) fm, \( a_{nn} = -2.37 \) fm, \( a_{pp} = -18.8 \) fm, \( r_{np} = 2.85 \) fm, \( r_{nn} = 2.75 \) fm, and \( r_{pp} = 2.75 \) fm [13] we find the averages \( \Delta_D = -1.5 \) MeV and \( g_D^2 = 1.0 \cdot 10^{-3} \) MeV\(^{-1}\).

With the parameters so determined, we turn now to possible predictions in low-energy nucleon-deuteron scattering. For simplicity we restrict ourselves to scattering below the deuteron break-up threshold \( \frac{3}{4} \), where the S-wave is dominant. There are two S-wave channels, corresponding to total spin \( J = 3/2 \) and \( J = 1/2 \). In the quartet only \( D \) contributes while in the doublet \( T \) also appears. The \( Nd \) scattering amplitude \( T_{Nd} \) from the same interactions is given by the diagrams in Fig. 3, which can be summed up by solving an integral equation in the quartet and a pair of coupled integral equations in the doublet channel.

\footnote{In contrast to an EFT with an “elementary” deuteron field, we can in principle extend our results above the break-up threshold (as long as the typical momentum remains much smaller than the pion mass) at the expense only of a greater numerical effort.}
The first diagram in the right-hand-side of Fig. 3 gives a contribution of order $\sim M R^2 / p^2 \sim a^2 / M r_0$. The remaining graphs mix different orders in $r_0 / a$, since they involve the dibaryon propagator; they contribute at order $\sim g^4 M^2 / p \Delta \sim (a^2 / M r_0)(1 + \varphi(r_0 / a) + \ldots)$. Other two-body contributions not included in Fig. 3 are suppressed by at least three powers of $r_0 / 2a$. For instance, $P$-wave interactions arise from a term in the Lagrangian with two derivatives and a coefficient of $O(1 / M^2 m_o)$; substitution of one of the dibaryon propagators by a $P$-wave interaction vertex would thus be suppressed by $\sim (r_0 / a)^3$ in comparison to the leading order. Likewise, the effect of the higher derivative term (not written explicitly in (1)) responsible for the shape parameter $\sim k^4$ in the effective range expansion of the nucleon-nucleon interaction is $\sim k^4 r_0^3$, and is thus also suppressed by $(r_0 / a)^3$ compared to the leading piece $\sim 1 / a$.

The diagrams in Fig. 3 are power-counting finite, but this does not preclude the existence of relevant contact interactions between the nucleon and the dibaryons. Pion exchange that would generate such interactions can be expected to be larger for the $I = 1$ dibaryon $T$, and therefore predominantly affect the $J = 1/2$ channel. Moreover, in the $J = 3/2$ channel, all spins are aligned and the three nucleons cannot be at the same position, so a six-nucleon vertex should contain at least two derivatives and have a coefficient $\sim 1 / M m_o^2$; its contribution is thus expected to be suppressed in relation to the leading order graph by six powers of $r_0 / a$. We will return to the doublet case in a future publication. Here we study the quartet channel, expected to be much less sensitive to the details of the physics of momenta of $O(m_o)$.

An enormous simplification comes about because the s-channel interaction due to the dibaryon is both local and separable. This allows us to write a simple integral equation that sums all the graphs in Fig. 3. Performing the integration over the time-component of the loop 4-momentum, we find that the conveniently normalized on-shell amplitude as a function of the initial (final) center-of-mass 3-momentum $k (p)$ satisfies

$$
\left[ -\frac{3(p^2 - k^2)}{8 M^2 g_\sigma^2} + \frac{1}{4 \pi} \sqrt{\frac{3}{4} (p^2 - k^2) + MB - \sqrt{MB}} \right] \frac{\psi(p, k)}{p^2 - k^2 - i \epsilon} = \frac{1}{(p + k/2)^2 + MB} - \int \frac{d^3 l}{(2 \pi)^3} \frac{1}{l^2 + l \cdot p + p^2 - \frac{3}{4} k^2 + MB} \frac{\psi(l, k)}{l^2 - k^2 - i \epsilon}.
$$

Note that all terms in a perturbative expansion of $\psi$ in (5) are of the same order ($\sim 1 / \sqrt{MB}$). It is straightforward but tedious to show that the wave function $(2 \pi)^3 \delta(p - k) + \psi(p, k) / (p^2 - k^2 - i \epsilon)$ corresponding to a scattering solution indeed satisfies the Schrödinger equation derived from the Lagrangian (1).
At zero energy \((k \to 0)\) the calculation simplifies: only the S-wave, dependent on the magnitudes of momenta, contributes to the scattering, and we can perform the angular integration directly. It is also convenient to normalize all quantities to \(\sqrt{MB}\). Defining \(x = p / \sqrt{MB}\),

\[
a(x) = \frac{\sqrt{MB}}{4\pi} \ln \left( \frac{p}{\sqrt{MB}} , 0 \right),
\]

and introducing

\[
F(x, z) = \frac{1}{xz} \ln \left( \frac{x^2 + z^2 + 1 + xz}{x^2 + z^2 + 1 - xz} \right),
\]

Eq. (5) becomes

\[
\frac{3}{4} \left[ -\eta + \frac{1}{1 + \sqrt{1 + \frac{3}{4} x^2}} \right] a(x) = -\frac{1}{x^2 + 1} - \frac{1}{\pi} \int_0^\infty dz F(x, z) a(z).
\]

Note that there is only one parameter \(\eta = 2\pi\sqrt{MB} / M^2 g_0^2 = 3 r_0\sqrt{MB} / 2 = 0.40\) in this equation. The value of the function \(a(x)\) at \(x = 0\) gives the \(Nd\) scattering length in units of \(1/\sqrt{MB}\). The same equation was previously obtained and solved in the zero-range limit (\(\eta \to 0\)) [15].

We have solved Eq. (8) numerically for \(\eta = 0.40\) by the Nystrom method [16]. The solution \(a(x)\) is plotted as the solid line in Fig. 4. The pole in \(a(x)\) around \(x \sim 4.4\) is associated with the spurious deep two-body pole. Its presence allows intermediate states where two nucleons fall into this deep state while the other has extra energy. The interesting point is that even though the effective theory makes nonsensical predictions outside its domain of validity, like the existence of this new state, the low-\(x\) part of the curve is insensitive to the large-\(x\) behavior, and the prediction for the scattering length is sensible. In order to demonstrate this more explicitly

---

5 Because we account for the range of the \(NN\) interaction, we are able to calculate the energy-dependence of \(Nd\) scattering close to threshold as well [14].

6 Since this deep pole is a ghost and does not appear in initial or final states, it should not cause problems in any process where typical momenta are within the range of validity of the EFT.
we have also solved Eq. (8) with a cut-off two-nucleon amplitude without the deep pole. For a cut-off of 150 MeV we obtain the broken line in Fig. 4.

The quartet scattering length is \( a = -a(0)/\sqrt{MB} \). For \( n = 0 \) and \( B \) fixed, we reproduce the result \( a = 5.09 \) fm of Ref. [15]. Taking into account the finite range \( \eta = 0.40 \) we obtain (Fig. 4) \( a = 6.33 \) fm with an uncertainty from higher orders of \( \pm 0.10 \) fm. This result obtained with no free parameters is in very good agreement with the experimental value of \( a = 6.35 \pm 0.02 \) fm [17].

We thank David Kaplan for extensive discussions. Discussions with Ben Bakker, Joe Carlson, Vitaly Efimov, Jim Friar, Walter Glöckle, and Martin Savage are also acknowledged. UvK is grateful to Justus Koch for hospitality at NIKHEF where part of this work was carried out. This research was supported in part by the DOE grants DOE-ER-40561 (PFB) and DE-FG03-97ER41014 (UvK).

References

Band termination in the shell model

Andrius Juodagalvis, Sven Åberg

Mathematical Physics, Lund Institute of Technology, P.O. Box 118, S-221 00 Lund, Sweden

Received 17 December 1997; revised 10 March 1998
Editor: J.-P. Blaizot

Abstract

The mechanism of band termination is discussed from the spherical shell model point of view, and the role of breaking nucleon pairs within a given \( j \)-shell configuration is stressed. Two main types of band terminations are discussed. In one type the band termination includes a sequence of states within an approximative seniority coupling scheme and therefore is smooth. In the other type the terminated state involves the breaking of one additional nucleon pair that results in a less smooth band termination. Band terminations in nuclei around \( {^{48}}\text{Cr} \) are calculated in the \( fp \)-shell model, and both types of band terminations are illustrated.

PACS: 21.10.-k; 21.10.Re; 21.60.-n; 21.60.Cs

Keywords: Shell model; Band-termination; Rotation; \( fp \)-shell nuclei

Rotational bands that terminate is an interesting and rather frequently occurring phenomenon in nuclear physics, see e.g. Refs. [1–3]. A rotational band may terminate when all the single-particle angular momenta that is contained in a few valence particles/holes is coupled to its maximum value. Higher angular momenta then require more drastic re-arrangements of nucleons, resulting in a disruption of the systematic behaviour in the rotational band of energy and electromagnetic properties. Terminating rotational bands have now been observed in several nuclei ranging from the region around \( {^{20}}\text{Ne} \), the neutron deficient \( A = 110 \) region, I and Xe isotopes with mass numbers around \( A = 120 \), to the region around \( {^{158}}\text{Er} \), see Ref. [3] for explicit references. More recently, terminating bands have been found also in nuclei with a few valence particles/holes outside \( {^{56}}\text{Ni} \) [4], and around \( {^{48}}\text{Cr} \) [5].

The mechanism of band termination is well understood from the deformed shell model point of view: Low spins are built from a collectively rotating (prolate) nucleus. With increasing spin more and more particles align their spin vectors along the rotation axis and the shape selfconsistently changes from the prolate shape over triaxial shapes, finally reaching an axially symmetric shape, which is usually oblate, and where the symmetry axis coincides with the rotation axis. This non-collective symmetric state then forms the terminated state.

In this Letter we aim at understanding the band-termination scenario from the spherical shell model point of view. Band termination in very light nuclei like \( {^{20}}\text{Ne} \) is well understood from the Elliott model.
[6], but in heavier nuclei a more complex situation emerges. We shall here focus mainly on nuclei in the lower part of the \( fp\)-shell. This region is especially suited for this study since the nuclei are heavy enough to show complex rotational behaviour, backbending and band terminations in a similar way as much heavier nuclei, and is still accessible for extensive shell model calculations.

By obtaining a detailed microscopic understanding of the band-termination phenomenon within the spherical shell model, and comparing the results with the deformed shell model (e.g. cranked Nilsson-Strutinsky), we may gain insight into what parts of the residual interaction are missing in the deformed models.

We show that consecutive decouplings of neutron and proton \( T = 1 \) pairs play an important role in building up angular momentum. This simple coupling scheme is valid when the collectivity is weak, as for example close to band termination, and sequences of spin states can then be approximately described by the \( j^2\)-coupling scheme (see inset in Fig. 2). A \( j^2\)-coupled spin sequence involves two alike nucleons and adds \( 0, 2, ..., 2\jmath - 1 \) \( h \) to the total angular momentum value with little influence from the rest of the nucleons. If the \( j^2\)-spin sequence includes the terminated state, the band termination becomes smooth and the terminated state is relatively favoured in energy. If, however, the terminated state is created by the decoupling of one additional \( T = 1 \) pair, i.e. by an increase in seniority, the band termination may become non-smooth and unfavoured due to the extra pairing energy that is lost in this process. Nuclei with a small number of valence particles/holes outside a spherical core are non-collective also at low spins, and traces of the proposed coupling scheme can then be seen throughout the yrast sequence of states.

Since we are studying nuclei with \( N = Z \) the neutron-proton interaction plays an important role, and deviations from the above simple picture emerges. This is analysed in terms of a simple, exactly solvable pairing model with neutrons and protons occupying the same \( j\)-shell [7]. For example, the non-smooth type of band termination (when seniority increases) may be more or less unfavoured depending on the relative seniority of protons and neutrons (described in terms of the "reduced isospin", \( \jmath \)). Due to the \( T = 1 \) pairing, the most unfavoured type of band termination thus appears when the increase in seniority makes the neutron and proton seniority more equal in the terminated state (\( \Delta \jmath = -1 \)).

The study is performed within the spherical shell model using the RITSSCHIL code [8] with the modified surface delta interaction [9],

\[
V_{MSD1}(r_1, r_2) = -4\pi A'_T \delta(r_1 - r_2) \delta(r_1 - R_0) + B'(\tau_1 \cdot \tau_2).
\]

(1)

The values of the parameters \( A'_T \) were adjusted to fit the yrast states in \(^{48}\text{Cr}\), while the \( B' \) value was fitted to the excitation energy of the yrast odd-spin band in \(^{46}\text{V}\). This resulted in \( A_1 = 0.67, A_0 = 0.35 \) and \( B = 0.37 \), where the values are multiplied by a constant factor coming from the radial integral. The single-particle energies used in the calculations are \( E_{0\jmath = 1/2} = 0.0 \text{ MeV} \) and \( E_{1\jmath = 1/2} = 2.1 \text{ MeV} \).

Let us first study the nucleus \(^{48}\text{Cr}\) which is one of the heaviest \( N = Z \) nuclei available for extensive measurements. In Fig. 1 we show some of the measured energy states (crosses) in \(^{48}\text{Cr}\). The spectrum exhibits rotational behaviour, backbending of the yrast band, as well as band termination. The experimental data are extremely well described by the Madrid-Strasbourg shell model calculations including the full \( fp\)-shell [10]. But also the present calculations, which are performed in the restricted shell model space of \( f_{3/2}p_{3/2} \) ("quasi SU(3)" [11,12]), describes the data very well. In Fig. 1 we show calculated energies (open circles) of the yrast sequence of states, the continuation of the ground-state band, and the lowest \( T = 1 \) band.

The ground-state band exhibits rotational behaviour with \( E(4^+) / E(2^+) \approx 2.5 \), and \( B(E2; 2^+ \rightarrow 0^+) = 31 \text{ W.u.} \) [13], corresponding to a deformation, \( \beta \approx 0.3 \), if the deformed rotor picture is adopted. At \( I = 10 \) the structure of the yrast band changes and a rather sharp backbending appears.

Above the backbending the yrast states are of rather pure \( f_{3/2} \) character, and the measured \( B(E2) \) values smoothly drop as the terminated state at \( 16^+ \) at \( E = 13.31 \text{ MeV} \) is approached [14], in nice agreement with the shell model calculation [10]. The terminated \( 16^+ \) state is mainly described by \( f_{3/2}^2 \). This state is quite unique in the sense that it has an almost spherical intrinsic shape, where the sphericity...
Fig. 1. Three rotational bands in $^{48}$Cr that terminate at $I^+ = 16^+_1$, $16^+_2$, and $15^+_2$, respectively. Crosses denote experimental data [5] and small open circles the present calculations. The yrast band as well as the continuation of the ground-state band have even spin values and isospin, $T_s = 0$, while the more excited band has odd spin values and isospin, $T_s = 1$. States that terminate a rotational band are encircled. In b) a rotational energy of 0.051 $\hbar I$ has been subtracted. The calculated terminated states at $16^+_1$ and $15^+_2$ constitute examples of non-smooth $D_{3/2}$ and smooth $D_{1/2}$ band terminations, respectively.

appears because all positive $m$-components in the \( f_{3/2} \) shell are occupied with a coherent coupling of the spin vectors, $I = 2(7/2 + 5/2 + 3/2 + 1/2) = 16$.

The ground-state band in $^{48}$Cr apparently continues above the backbending, although the structure of the states changes. In the present $f_{7/2} p_{3/2}$ calculation the configuration of the states above the backbending is dominated by one particle being excited to the $p_{3/2}$ shell, $f_{7/2} p_{3/2}$, and the band terminates at the same angular momentum as the yrast band, $16^+$.

The lowest $T = 1$ states with odd angular momentum values form a rotational band which terminates at $15^+_2$ (calculated at $E = 14.2$ MeV) as a predominately $f_{7/2}^0$ state, i.e. a $T = 1$ partner to the yrast $16^+$ state.

An observed $16^+_2$ state at $E = 15.73$ MeV, that mainly decays to the $14^+_1$ state, may correspond to the calculated terminated $f_{7/2}^1 p_{3/2}^1$ state. No states belonging to the suggested $T = 1$ band in $^{48}$Cr have been identified, but the yrast states in $^{48}$V (with odd spins) constitute isobaric analog states and have been observed [15]. This band will be discussed below.

The yrast band of $^{48}$Cr has been calculated also in the cranked Nilsson-Strutinsky model (CNS) [16]. These calculations, which do not include the pairing force, can quite well describe the general behaviour above the backbending. The band termination is seen as a gradual change of deformation from a prolate shape with the deformation $\varepsilon = 0.23$ at $I = 0$ to an almost spherical terminated state. The experimental energies above the backbending are rather well described, except for the terminated state that in the CNS model follows the continuation of the ground-state band energies, while the data (see Fig. 1b) show that it is unfavoured by about 1 MeV. Residual interactions outside the CNS model (particularly $T_s = 1$ pairing) are thus responsible for this energy shift, that is nicely reproduced in the present spherical shell model calculation.

On the other hand, the $T = 1$ band, which is also predominantly built from eight particles in $f_{7/2}$, terminates in a quite smooth way. We thus have quite different types of behaviour in the band termination of the $T = 0$ and $T = 1$ “$f_{7/2}$” bands. While the terminated state in the $T = 0$ band is unfavoured with about 1 MeV as compared to a linear extrapolation, the terminated state in the $T = 1$ band is slightly favoured, see Fig. 1b. We shall below explain these two different kinds of band terminations.

In order to better understand the different behaviour of band terminations, we perform a shell model calculation of the yrast spectrum of $^{48}$Cr where the available shell model states are restricted to the $f_{7/2}$ shell. In Fig. 2 the result of this restricted calculation is compared with the $f_{1/2} p_{3/2}$ calculation which resembles the experimental data, as discussed above. Low spins are quite poorly described in the $f_{1/2}$ calculation. The states above the back-
is constructed by first minimizing the seniority, $T$, that for given isospin, $I$, and $T$ protons and $n$ neutrons in a restricted model space of $f$.

This is, of course, due to the rather pure bending are better described, and the $16^+_1$ state is quite well described in the restricted model space calculation. This is, of course, due to the rather pure $(f_{7/2})^6$ character of this state.

By an additional simplification, assuming a pure pairing force ($\delta_{I_\pi=0}$), an analytical expression can be given for the pairing energy of the system with $n_p$ protons and $n_n$ neutrons in a $j$-shell [7],

$$P(n,T,v,t) = \frac{n-v}{4}(4j + 8 - n - v) - T(T + 1) + t(t + 1),$$

(2)

where $n = n_p + n_n$ is the total number of particles, $v = v_p + v_n$ is the total number of unpaired particles (seniority), $t = \frac{1}{2} |v_p - v_n|$ is the reduced isospin, and $T$ is the total isospin. We notice from Eq. (2) that for given isospin, $T$, the yrast sequence of states is constructed by first minimizing the seniority, $v$, and subsequently maximizing the reduced isospin, $t$: i.e. breaking proton and neutron pairs as asymmetrically as possible. This simple expression for the pairing energy turns out to be a quite good approximation of the $f_{7/2}$ calculation adopting the modified surface delta interaction (Fig. 2), and gives further insight into our understanding of the observed spectra. Quantum numbers $(v,t)$ from Eq. (2) are shown in Fig. 2 for each state as obtained from the $f_{7/2}$ calculation. Highest angular momentum states for given seniority are marked out by squares.

Despite large deviations from the measured data in the restricted model-space calculation, it is still of interest to note the character of the states along the yrast line in the $f_{7/2}$ calculation, and their classification in terms of seniority and reduced isospin. The $T = 1 (f_{7/2})^2$ seniority coupling (see inset in Fig. 2) creates the arc-behaviour from angular momentum 0 to 6. In the $f_{7/2}$ calculation the $6^+$ state is thus described as a linear combination of one aligned proton pair, $(\pi f_{7/2}^2)(\pi f_{7/2}^2)$, and one aligned neutron pair, $(\nu f_{7/2}^2)(\pi f_{7/2}^2)$, and has $(v,t) = (2,1)$. The $8^+$ state is obtained by aligning two pairs of one kind of particles (4,2), and by the alignment of one neutron pair and one proton pair (4,0), the $10^+$ and $12^+$ states are built. The $12^+$ state is thus described as four maximally aligned particles (two protons and two neutrons), $I = 2(7/2 + 5/2) = 12$ with $(v,t) = (4,0)$, while the other four particles are pairwise coupled to $I = 0$. Angular momenta 14 and 16 are obtained by a subsequent breaking of these two pairs, i.e. $(v,t) = (6,1)$ and $(8,0)$, respectively. The band termination is thus connected with an increase in seniority, $\Delta v = 2$, implying an extra energy cost. In addition, the reduced isospin decreases, $\Delta t = -1$, implying an additional energy cost for the terminated state, see Eq. (2).

It is interesting to note that this maximally unfavoured way of band termination, $\Delta v = 2$, $\Delta t = -1$, only appears for exactly half-filled $j$-shells with equal numbers of protons and neutrons. For example, the expected band termination at $25^+$ in $^{90}_{45}$Rh$_{45}$, with 5 protons and 5 neutrons in $g_{9/2}$, is thus expected at a considerably higher energy (1-2 MeV) than the continuation of the yrast sequence from ...$19^+$, $21^+$ and $23^+$, when the energy of the different states are compared after the rotational reference has been subtracted, cf. Fig. 1b.

The corresponding situation in the $sd$-shell appears for the ground-state band of $^{20}_{11}$Na$_{11}$ with three protons and three neutrons in $d_{5/2}$, i.e. a half-filled $j$-shell. The sequence of odd-spin members, $3^+$, $5^+$, $7^+$ and $9^+$, of the ground-state band $(K^\pi,T) = (3^+,0)$ is thus expected to terminate at $9^+$ with the terminated state about 1 MeV higher in energy than a smooth continuation of the lower spin states would.

{Fig. 2. The yrast rotational band in $^{48}$Cr is calculated in the restricted model space of $(f_{7/2})$ (dashed line) and the $(f_{7/2})^2$ space (solid line), using the same set of parameters for the modified surface delta interaction. States where one, two, three or four pairs of nucleons are coupled to give maximum angular momentum in the $(f_{7/2})$ space, as allowed by the Pauli principle, are marked by squares. Seniority and reduced isospin, $(v,t)$, from Eq. (2) are shown for each state in the $(f_{7/2})$ calculation. The behaviour of one single pair of alike nucleons in $(f_{7/2})^2$ ($f^2$-coupling scheme) is shown in the inset.}
suggested. And indeed, this is exactly what is seen experimentally [17]. In the cranked Nilsson-Strutinsky model the states of the band are described [18] as smoothly changing the deformation from a collective, prolate shape with $\varepsilon = 0.40, \gamma = 0^\circ$ at $I^\pi = 3^+$ to a band termination at $I^\pi = 9^+$ with $\varepsilon = 0.09, \gamma = 60^\circ$. As in the case of $^{48}$Cr, discussed above, the NS-calculations predict the terminated state too low in energy (by about 0.7 MeV).

Yet another example of a $\Delta \nu = 2, \Delta \ell = -1$ band termination is the positive-parity band in $^{46}$Cr that has one neutron hole in $d_{3/2}$ and 4 protons and 4 neutrons in $f_{3/2}$ (as for the discussed states of $^{48}$Cr). The band termination at $35^+_2$ is thus also expected to be particularly unfavoured.

Obviously, the simple picture from the pure $j$-shell discussion is far from being correct for $^{48}$Cr, at least for the lower spins. When the energy distance from the $f_{7/2}$ shell to the $p_{3/2}$ shell decreases (this distance can be considered as infinitely large in the pure $f_{7/2}$ calculation), the lower spin states pick up a larger and larger fraction of $p_{3/2}$, and the arc-behaviour between $I = 0–6$ (and partly between $I = 6–12$) smoothly changes into a rotational behaviour, thus smearing out the irregularity at $I = 6$. This is seen in Fig. 2 where the rotational behaviour of the ground-state band is clearly seen in the $f_{7/2}p_{3/2}$ calculation with the distance to the $p_{3/2}$ shell being 2.1 MeV. The mixing between $f_{7/2}$ and $p_{3/2}$ is then so strong that no traces remain of the arc-behaviour and the aligned character of the $6^+$ state.

This onset of rotation (deformation) through the increased mixing of $p_{3/2}$ in the shell model calculations described above, reminds of the onset of deformation in the Nilsson model, where the single-particle wave functions describing the $\Omega = 1/2$ and $3/2$ levels emerging from the $f_{7/2}$ shell ($330 \ 1/2$) and $[321 \ 3/2]$ get a larger and larger mixing of $p_{3/2}$ (which have the same $O$-components) as the deformation parameter is increased.

The mixing with the $p_{3/2}$ shell is less important for nuclei with a smaller number of valence particles outside the $N = Z = 20$ core. It is also less important for angular momentum states close to band termination, even for nuclei which show a clear rotational behaviour at low spins, as for $^{48}$Cr. One then expects that the arc-behaviour, or traces of this coupling scheme, should be visible in (parts of) the sequence of angular momentum states. In Fig. 3 the yrast sequences of angular momentum states are shown for six $fp$-shell nuclei and one $sd$-shell nucleus. Experimental data are shown when available, else the calculated spectrum is shown. We have performed detailed calculations for all nuclei shown in Fig. 3, and the agreement with the measured spectra is generally very good. States with maximal angular momentum for given seniority (in the single-$j$ calcu-

![Fig. 3. a) Examples of smooth band terminations ($\Delta \nu = \Delta \ell = 0$), see Eq. (2): yrast bands in $^{20}$Ne, $^{44}$Ti, $^{46}$V, and the odd spin members of $^{48}$V. In b) three examples of non-smooth band terminations ($\Delta \nu = 2$) are shown: yrast bands in $^{46}$Ti, $^{48}$Cr, and lowest signature states in $^{50}$Cr ($\Delta \ell = -1$). Measured states are shown by crosses and calculated states by open circles. States where a pair of nucleons (in the calculations with the model space restricted to a single $j$-shell) is coupled to give maximum angular momentum are marked by squares, cf. Fig. 2. To facilitate the reading of the figure, 1 MeV has been added to all energy states in $^{46}$V and $^{48}$Cr. Experimental data are taken from Refs. [19] (Ne), [20] (Ti), [21] (V), [15] (V), [5] (Cr) and [22] (Cr).]
lution) are marked by squares, and the arc-behaviour
is in many cases built between these states. In Fig.
3a the bands end with one square \((\Delta \nu = 0\) ter-
mination, smooth) while in Fig. 3b the bands end with
two squares \((\Delta \nu = 2\) termination, non-smooth).

The nucleus \(^{44}\text{Ti}\) with only four valence particles in
the \(fp\)-shell clearly shows the expected arc-behav-
ior from \(6^+\) to \(12^+\) \((\nu,t) = (4,0)\) for \(I = 8, 10, 12\).
The band terminates in a smooth way at \(12^+\) since
\(I = 12\) constitutes the highest spin state in the
\(j^2\)-coupling scheme involving the second \(T = 1\) pair.
In terms of seniority and reduced isospin (Eq. (2))
either of these quantum numbers change for the
terminating state, i.e. \(\Delta \nu = \Delta t = 0\), implying a
comparably low energy cost to reach the terminated
state.

The sd-shell version of \(^{44}\text{Ti}\) is obviously \(^{20}\text{Ne}\) with
four valence particles in the sd-shell. Since two
particles in \(d_{5/2}\) couple to a maximal spin of \(4\hbar\) the arc-
bows now have a period of \(4\hbar\). This implies the
\(4^+\) state is slightly below the rotational band ener-
gies, and the \(8^+\) state is a strongly favoured termina-
ted state (\(\Delta \nu = \Delta t = 4\) staggering').

The odd-spin members of the observed ground-
state band in \(^{48}\text{V}\) have \(T = 1\) and constitute isobaric
analog states with the calculated lowest \(T = 1\) band in
\(^{48}\text{Cr}\) (Fig. 1). The two isobaric analog bands are
identical in the calculations, and we here choose to
discuss the observed band in \(^{48}\text{V}\), see Fig. 3a. As
can be seen in Fig. 1b (\(T = 1\) band in \(^{48}\text{Cr}\)) the band
terminates in a smooth way with the \(15^+\) state. In a
similar way as for \(^{44}\text{Ti}\) the terminated state in \(^{48}\text{V}\) is
the last state in the angular momentum coupling of
two particles (i.e. \(\Delta \nu = \Delta t = 0\)) explaining its
smoothness. However, since the mixing with \(p_{3/2}\) is
larger in \(^{48}\text{V}\) than in \(^{44}\text{Ti}\) the spin sequence just
before band termination in \(^{48}\text{V}\) is more rotational
like than in \(^{44}\text{Ti}\), see Fig. 3a.

By removing one pair of neutrons from \(^{48}\text{V}\), and
thus reaching \(^{46}\text{V}\), the coupling scheme does not
change since one pair in \(^{48}\text{V}\) is all the time coupled
to spin zero \((m = \pm 1/2)\). The only difference is
that due to less particles in \(^{46}\text{V}\) the rotational
behaviour is suppressed, and the energy sequences
become more arc-like.

The non-smooth band termination observed in the
yrast band of \(^{48}\text{Cr}\) was above explained as the
consequent breaking of the two remaining pairs above
\(I = 12\), giving \(I = 14\) and 16, respectively. If one of
the two pairs of nucleons is removed from \(^{48}\text{Cr}\) we
get \(^{46}\text{Ti}\) (or \(^{46}\text{Cr}\)). Consequently, for this nucleus the
last angular momentum state in \(^{48}\text{Cr}\) \((16^+)\) is not
available, and the termination appears at \(I^\rho = 14^+\)
instead. This band termination is connected with an
increase of the seniority, \(\Delta \nu = 2\), that causes an
extra energy cost for the terminated state. However,
since the reduced isospin increases for the terminated
state, \(\Delta t = 1\), the energy cost for the band termina-
tion is considerably smaller than for \(^{48}\text{Cr}\) where
\(\Delta t = -1\). Another effect of the removal of the pair
is that the rotational behaviour seen in \(^{48}\text{Cr}\) is now
changed, and traces of an arc-behaviour in \(^{46}\text{Ti}\) is
seen between \(I = 0-6\) (cf. Fig. 2), see Fig. 3b.

Another example of a non-smooth band termina-
tion is seen in \(^{49}\text{Cr}\), see Fig. 3b. \(I = 27/2\) is the
highest angular momentum that can be obtained with
seniority \(\nu = 5, (\nu,t) = (5,1/2)\). A maximum of addi-
tionally \(2\hbar\) is obtained by breaking the remaining
proton pair in the \(f_{7/2}\) shell. This process costs some
extra energy \((\Delta \nu = 2\) and \(\Delta t = 0\)\) and the band
termination at \(31/2^+\), \(7,1/2\), is non-smooth (and
somewhat unfavoured).

In summary, we have shown that the seniority-
coupling scheme can be schematically used in order
to obtain a microscopic understanding of the band-
termination phenomena in nuclei around \(^{48}\text{Cr}\). This
picture implies two kinds of band terminations,
smooth and non-smooth, depending on whether the
terminated state belongs to the \(j^2\)-coupling sequence
\((\Delta \nu = 0\) or not \((\Delta \nu = 2)\). In the latter case the
reduced isospin may change in three different ways,
\(\Delta t = -1, 0\) or \(+1\), corresponding to an additional,
no additional, or a reduced increase in the energy for
the terminated state. The onset of collectivity (defor-
mation) emerges due to an increased mixing of the
configurations, and a smooth transition from the
\(j^2\)-coupling scheme to the rotational picture was
found. We could then understand also the yrast
sequence of states in several \(fp\)-shell nuclei, and
observed irregularities in `"rotational"' bands could
be explained from this picture. Comparisons with
CNS calculations suggested that the \(T = 1\) pairing
(particularly the np-part) may play an important role
for \(N \approx Z\) nuclei also around band termination.

We thank Silvia M. Lenzi for several discussions
and for allowing us to use un-published data. Com-
ments on the manuscript from Ingemar Ragnarsson are acknowledged. One of the authors (A.J.) thanks Magdalena Gorska for help in setting up the RITS-SCHIL computer code. We also thank the referee who drew our attention to Eq. (2). S.A. thanks the Swedish National Research Council (NFR) for financial support.

References

Radiative energy-loss of heavy quarks in a quark-gluon plasma

Munshi Golam Mustafa a, Dipali Pal b, Dinesh Kumar Srivastava b, Markus Thoma c

a Saha Institute of Nuclear Physics, 1 / AF Bidhan Nagar, Calcutta 700 064, India
b Variable Energy Cyclotron Centre, 1 / AF Bidhan Nagar, Calcutta 700 064, India
c Institut für Theoretische Physik, Universität Giessen, D-35392 Giessen, Germany

Received 1 December 1997; revised 14 March 1998

Abstract

We estimate the radiative energy-loss of heavy quarks, produced from the initial fusion of partons, while propagating in a quark-gluon plasma which may be formed in the wake of relativistic heavy ion collisions. We find that the radiative energy-loss for heavy quarks is larger than the collisional energy-loss leading to a complete stopping of initially produced quarks with energies below about 10 GeV within 2 fm. We point out the consequences on possible signals of the quark-gluon plasma.

One of the most interesting predictions of QCD is the transition from the confined/chirally broken phase to the deconfined/chirally symmetric state of quasi-free quarks and gluons, the so-called quark-gluon plasma (QGP). Relativistic heavy ion collisions are being studied with the intention of investigating the properties of the QGP [1]. While experiments at AGS and SPS continue, new experiments have been planned at RHIC and LHC with centre of mass energies 200 AGeV and 5.5 ATeV, respectively. During the past decade many different signatures of the transition to the QGP have been proposed. A promising example is the emission of penetrating probes such as dileptons and single photons, which can reveal the early parton dynamics and the history of evolution of the plasma. Similarly, the production and propagation of open charm and high energy jets in a dense medium, can provide information [2] about parton scattering and thermalization of the partonic system. Jets are expected to show up at collider energies at RHIC and LHC.

Heavy quark pairs are mostly produced from the initial fusion of partons mostly from \( gg \rightarrow Q\bar{Q} \), but also from \( q\bar{q} \rightarrow Q\bar{Q} \), where \( q \) denotes one of the lighter quarks and \( Q \) is a heavy quark) of the colliding nucleons and also from the QGP, if the initial temperature is high enough, which is likely to be achieved at RHIC and LHC energies. The charm quarks will be produced on a time scale of \( 1/2m_c = 0.07 \text{ fm}/c \), which would be as low as \( \approx 0.02 \text{ fm}/c \) for bottom quarks. There is no production of heavy quarks at late times in the QGP and none in the hadronic matter. Thus, the total number of heavy quarks gets frozen very early in the history of the collision which makes them a good candidate for a probe of the QGP. Immediately upon their production, these heavy quarks will propagate through the deconfined matter and start losing energy. The en-
ergy-loss suffered by these quarks will determine the shape of the dilepton spectra produced from correlated charm (or bottom) decay which provides a large background to dilepton production from annihilation of quarks in the plasma. We shall come back to this aspect towards the end of this Letter.

There are two contributions to the energy-loss of a heavy quark in the QGP: one caused by elastic collisions with the light partons of the QGP and the other by radiation of the decelerated color charge, i.e., bremsstrahlung of gluons. There is an extensive body of literature [3–8] on the collisional energy-loss of energetic quarks considering elastic collisions with the quarks and gluons ($Qg \to Qg$ and $Qq \to Qq$) of the dense medium. A complete leading order result for the collisional energy-loss of heavy quarks has been found using the hard thermal loop resummation technique [9].

It is well known that the contribution of the radiative processes ($Qq \to Qqg$ and $Qg \to Qgg$) is of the same order in the coupling constant as the collisional energy-loss [9]. The estimate of the radiative energy-loss in the past has been discussed by a number of authors [10–12] within perturbative QCD taking into account the Landau-Pomeranchuk suppression due to multiple collisions. These studies, however, were limited to the case of massless energetic quarks and gluons. As far as we know, there is no estimate of the radiative energy-loss for heavy quarks in the literature. It is not easy to extend the sophisticated treatment of multiple scattering formulated by the authors of Ref. [12] to the case of heavy quarks.

Until such a detailed investigation is performed, an extension of the work of Ref. [10] for the radiative energy-loss of massless quarks to the case of heavy quarks can provide valuable insight. This approach is similar to the one by Gyulassy, Wang and Plümper [11] and leads to almost identical results in the case of light partons. Baier et al. [12], on the other hand, found a different dependence of the radiative energy-loss on the energy of the parton by including the rescattering of the emitted gluon in the QGP. However, inserting typical values for the energy of the parton, the temperature, and the coupling constant yields quantitatively similar results. Therefore we will restrict ourselves to a simple approach similar to the one of Ref. [10] in the present work and discuss the consequences of our estimate for signatures of the QGP.

We start from an expression for the gluon emission probability which has been derived by Gunion and Bertsch [13] in the case of light partons assuming a factorization of the matrix elements of Fig. 1 into elastic scattering and gluon emission. It can be shown that their result also holds if one of the light quarks in Fig. 1 is replaced by a heavy quark assuming that the energy of the emitted gluon is not too large ($q_0 \ll \sqrt{s}$). The final result for the multiplicity distribution of the radiated gluon obtained in light cone gauge, where also the gluon radiation from the

---

Fig. 1. Feynman diagrams for gluon bremsstrahlung from quarks.
lower quark lines in Fig. 1 has been taken into account, can then be written as
\[
\frac{dn_g}{d\eta d^2q_\perp} = \frac{C_A\alpha_s}{\pi^3} q_\perp^2 \left( q_\perp^2 - l_\perp^2 \right)^{-1}.
\] (1)

where \( q = (q_0, q_\perp, q_\parallel) \) and \( l = (l_0, l_\perp, l_\parallel) \) are the four momenta of the emitted and the exchanged gluons, respectively, and \( \eta = (1/2)\ln[(q_0 + q_\parallel)/(q_0 - q_\parallel)] \) is the rapidity. \( C_A = 3 \) is the Casimir invariant of the adjoint representation. The factorization in (1) was obtained in the limit \( \lambda l_\perp \ll q_\perp \), where \( \lambda \) is the fractional momentum carried by the radiated gluon relative to the maximum available, such that the radiation is confined to a uniform (central) rapidity region. Eq. (1) holds also for gluon bremsstrahlung emitted by a gluon.

Now we can estimate the radiative energy-loss per unit length for heavy quarks by multiplying the interaction rate \( \Gamma \) and the average energy-loss per collision \( \nu \), which is given by the average of the probability of radiating a gluon times the energy of the gluon. One can further correct for the Landau-Pomeranchuk effect by including a formation time restriction [10] through a step function \( \theta(\tau - \tau_f) \). This puts a restriction on the phase space of the emitted gluons in which the formation time, \( \tau \), must be smaller than the interaction time, \( \tau = 1/\Gamma \). The formation time is estimated by requiring the separation between the emitted gluon and the parton from which it is emitted to be \( r_\perp = \nu l_\perp \) or \( q_\perp \equiv |q_\perp| \) according to the uncertainty principle. Using \( \nu = q_\perp/q_0 \) and \( q_0 = q_\perp \cosh\eta \), we find \( \tau_f = \cosh\eta/q_\perp \).

The average radiative energy-loss per collision is calculated as
\[
\nu = \langle n_g q_0 \rangle = \int d\eta d^2q_\perp \frac{dn_g}{d\eta d^2q_\perp} q_0 \theta(\tau - \tau_f).
\] (2)

The formation time restriction serves as an infrared cut-off for the integration over the rapidity \( \eta \). Performing the integrations in (2) in the limit \( q_\perp \tau \gg 1 \) and \( q_\perp \gg l_\perp \), we get

\[
\nu = \frac{6\alpha_s}{\pi} \langle l_\perp^2 \rangle \tau \ln \left( \frac{q_{\perp\text{max}}}{q_{\perp\text{min}}} \right).
\] (3)

As a consequence of the limit \( q_\perp \tau \gg 1 \), i.e., assuming that the mean free path \( \tau \cdot c \) of the heavy quark is large compared to the formation length \( r_\perp \), the energy loss per collision is proportional to \( \tau = 1/\Gamma \). Therefore the interaction rate \( \Gamma \) drops out of the radiative energy loss given by \( \Gamma \cdot \nu \).

For the infrared cut-off \( q_{\perp\text{min}} \) we choose the Debye screening mass of a pure gluon gas,
\[
q_{\perp\text{min}} = \mu_D = \sqrt{4\pi\alpha_s} T,
\] (4)

where \( T \) is the temperature of the system. This choice comes from using the Debye screening mass in the gluon propagator entering into (1) [11]. For heavy quarks of mass \( M \), the square of the maximum transverse momentum of the emitted gluon is given by
\[
\langle q_{\perp \text{max}}^2 \rangle = \frac{\left( s - M^2 \right)^2}{4s},
\] (5)

where \( s \) is the Mandelstam variable. To evaluate (5) we need to compute \( \langle s \rangle \) and \( \langle 1/s \rangle \) leading to
\[
\langle s \rangle = M^2 + 2p^2 E, \quad \langle 1/s \rangle = \frac{1}{4p^2} \ln \left[ \frac{M^2 + 2Ep^2 + 2pp^2}{M^2 + 2Ep^2 - 2pp^2} \right],
\] (6)

where \( E \) and \( p \) are the energy and momentum of a incoming heavy quark and \( p^i \) is the average momentum of the light quark or gluon of the QGP. The average value of \( p^i \) can be taken as \( \sim 3T \). Now, (5) becomes,
\[
\langle q_{\perp \text{max}}^2 \rangle = \frac{3ET}{2} - \frac{M^2}{4} + \frac{M^4}{48pT} \ln \left[ \frac{M^2 + 6ET + 6pT}{M^2 + 6ET - 6pT} \right]
\] (7)

The average momentum transfer of the scattering process is defined as
\[
\langle t^2 \rangle = \langle t \rangle = \frac{\int_0^{q_{\perp\text{max}}^2} dt \, d\sigma/dt}{\int_0^{q_{\perp\text{max}}^2} d\sigma/dt},
\] (8)
where the differential cross section for elastic scattering is
\[
\frac{d\sigma}{dt} = \frac{1}{16\pi(s-M^2)^2} |\mathcal{M}|^2, \tag{9}
\]
and the square of the matrix element, $|\mathcal{M}|^2$ can be obtained from Ref. [14]. In the limit $|t| < s$, the differential cross section in (9) can be approximated by
\[
\frac{d\sigma}{dt} \sim \frac{1}{t^2}. \tag{10}
\]
The cross section (10) leads to a quadratic infrared singularity integrating over $t$. This situation can be improved by using a hard thermal loop resummed gluon propagator [15], which takes into account Debye screening as well as dynamical magnetic screening. In this way the quadratic singularity is reduced to a logarithmic one. Assuming a cut-off for the remaining divergence, e.g. provided by a non-perturbative magnetic screening mass, Debye screening and magnetic screening are of the same order neglecting logarithmic corrections. Therefore we use the widely adopted approximation (see e.g. Ref. [3]) of choosing the Debye mass as a single infrared cut-off for the longitudinal and transverse gluon exchange.

We have checked that the modification of the energy-loss using the full expression for $|\mathcal{M}|^2$ is negligible (see below). Combining (8) to (10) we get
\[
\langle I_\perp^2 \rangle = \frac{\mu_D^2(q_{\text{max}}^\perp)^2}{(q_{\text{max}}^\perp)^2 - \mu_D^2} \ln \left( \frac{(q_{\text{max}}^\perp)^2}{\mu_D^2} \right). \tag{11}
\]
The radiative energy-loss for heavy quarks is then obtained by combining (3) and (11) and multiplying by $\Gamma = 1/\tau$,
\[
\left( -\frac{dE}{dx} \right)_{\text{rad}} = \frac{3\alpha_s}{\pi} \frac{\mu_D^2(q_{\text{max}}^\perp)^2}{(q_{\text{max}}^\perp)^2 - \mu_D^2} \ln \left( \frac{(q_{\text{max}}^\perp)^2}{\mu_D^2} \right). \tag{12}
\]
This result is based on the approximation $q_{\text{max}}^\perp \gg \mu_D$. However, in the case of the logarithmic approximation used here, the requirement $q_{\text{max}}^\perp > \mu_D$ is sufficient in practice [16].

Since the mass of the quark in this expressions enters only via the maximum transverse momentum (7) the radiative energy-loss of a heavy quark differs from the one of a massless quark only for small energies of the order of $M$. Let us also recall the expression for the collisional energy-loss of heavy quarks considered in Ref. [9] using the hard thermal loop resummation technique. In the domain $E << M^2/T$, it reads
\[
\left( -\frac{dE}{dx} \right)_{\text{coll}} = \frac{8\pi \alpha_s^2 T^2}{3} \left( 1 + \frac{n_f}{6} \right) \times \left[ \frac{1}{v} - \frac{1 - v^2}{2 v^2} \ln \left( \frac{1 + v}{1 - v} \right) \right] 
\times \ln \left[ \frac{n_f}{2^{n_f}} B(v) \frac{E T}{m_g M} \right], \tag{13}
\]
whereas for $E >> M^2/T$, it is
\[
\left( -\frac{dE}{dx} \right)_{\text{coll}} = \frac{8\pi \alpha_s^2 T^2}{3} \left( 1 + \frac{n_f}{6} \right) \times \ln \left[ \frac{n_f}{2^{n_f}} B(v) \frac{0.92 \sqrt{E T}}{m_g} \right]. \tag{14}
\]
where $v$ is the velocity of the heavy quarks, $B(v)$ is a smooth function of $v$, which can be taken approximately as 0.7, $n_f$ is the number of light quark flavours taken as 2.5, and $m_g = \gamma(1 + n_f/6)/3 g_T$ the thermal gluon mass.

It should be noted that the collisional and the radiative energy-loss are of the same order in the coupling constant [9], although the latter is caused by higher order diagrams within naive perturbation theory. In the absence of the Landau-Pomeranchuk effect the reason for this behaviour is the fact that the interaction rate entering into the radiative energy-loss suffers from a quadratically infrared singularity using a bare propagator for the exchanged gluon, whereas this divergence is reduced to a logarithmic one for the collisional energy-loss. This reduction is caused by the presence of the energy transfer of the exchanged gluon in the definition of the collisional energy-loss [9]. In the case of the radiative energy-loss, on the other hand, this factor is
The expression for the radiative energy-loss in (12) exhibits a threshold behaviour: for \((q^\perp)^2 < \mu_D^2\) there is no radiative energy-loss which is shown by the hatched area in Fig. 2. (Obviously, this behaviour will be different for mass-less quarks.) We see that the value of the threshold momenta of the heavy quarks, below which there is no radiation, increases with increasing temperature.

In Fig. 3, we compare our results with that of the collisional energy-loss for heavy quarks obtained in Ref. [9] as a function of energy at a temperature \(T = 500\) MeV and \(\alpha_s = 0.3\) for charm and bottom quarks. Before discussing our results, we give the justification for the approximation made in (10) for the differential cross section used in computing the radiative energy-loss, which enabled us to obtain the closed form given above. The solid lines represent the radiative energy-loss of heavy quarks with full

\[ G; a \]
\[ s \]

Taking account of the Landau-Pomeranchuk effect using the formation time restriction in the limit \(q^\perp \gg 1\) the interaction rate \(\Gamma\) does not appear in the radiative energy loss. However, it is replaced in \(\nu\) by the average momentum transfer which is of the same order neglecting logarithmic corrections: \(\langle I^2_f \rangle \sim \mu_D^2 \sim \alpha_s\).
$\frac{dE}{ds}$ whereas the dashed lines correspond to that with the approximate expression. We see that retaining only the $\sim 1/r^2$ term in $d\sigma/dt$ is sufficient for our purpose. The dash-dotted lines in Fig. 3 represent the collisional energy-loss obtained in Ref. [9]. We find that the radiative energy-loss dominates over the collisional one at all energies. For $E > 20$ GeV the difference amounts to an order of magnitude. Quarks with an energy below about 10 GeV are stopped already after about 2 fm. Therefore most of the initially produced quarks cannot escape from the fireball if a QGP phase is formed in a relativistic heavy ion collision.

The QGP expected to be produced at RHIC and LHC is likely to be far from chemical equilibrium, initially. Chemical reactions among the partons will then push it towards a chemical equilibrium [17–19]. The evolution of the temperature and the quark and gluon fugacities $f$ at RHIC and LHC energies has recently been obtained [19] for such a scenario with initial conditions from a Self Screened Parton Cascade Model [20]. As a first estimate, the expressions for the collisional energy-loss given earlier (Eqs. (13), (14)) can be modified by replacing the terms $(1 + n_l/6)$ by $(\lambda_q + \lambda_q n_l/6)$ and $n_l/(6 + n_l)$ by $\lambda_q n_l/(6 \lambda_q + \lambda_q n_l)$ to account for the departure from chemical equilibrium. Alternatively one may use the results of Ref. [21] for a non-equilibrium plasma.

The radiative energy-loss is modified by using the non-equilibrium Debye mass [18]

$$\mu_d^2 = 4\pi \alpha_x T^2$$

in (12).

This has interesting consequences. It has been shown recently [21] that considering only the collisional energy-loss, in this manner, amounts to having only a small drag on the motion of heavy quarks in such a plasma, at least at RHIC energies, where the charm quarks were found to loose only $\sim 10\%$ of their energy during their propagation, through the plasma and up to $40\%$ of their initial energy at LHC, in a collision involving two gold nuclei [22].

The drag acting on the heavy quark is conveniently defined by writing

$$- \frac{dE}{ds} = Ap,$$

where $A$ denotes a drag-coefficient in the spirit of the treatment used earlier in literature [4,21] and $p$ is the momentum of the heavy quark. Adding the collisional and radiative energy-loss experienced by a heavy quark in such an equilibrating and cooling plasma, we have found that

$$A = C/\tau,$$

where $C$ is a slowly varying function of $p$ with $C = 0.4$ for charm quarks at RHIC energies and $\sim 0.7$ at LHC energies, for $E \leq 5 - 6$ GeV. (The weak dependence of $A$ on $p$ is a consequence of the complicated momentum dependence of (12). From Fig. 3 we observe that the radiative energy loss of a charm quark can be assumed to depend linearly on $p$ for our purpose. This observation holds in particular for the energy range $E < 10$ GeV in which we are interested.)

This leads to a rather large drag of $\sim 1.6/fm$ at RHIC and $\sim 2.7/fm$ at LHC on charm quarks at $\tau = \tau_r$, where $\tau_r = 0.25$ fm/c is determined by the onset of the kinetic equilibrium [19].

This has a very important implication. Consider a charm quark having an energy $E_i$ of the order of a few GeV at time $\tau_i$ in such an expanding and chemically equilibrating plasma. Due to this large drag, the charm quark produced initially will come to rest very quickly and diffuse. We have verified this by performing numerical calculations of the final momentum of charm quarks which propagate under such a drag. We find that, irrespectively of the initial energy ($E_i \leq 5–6$ GeV) the final energy of the charm quark is about $1.5 - 1.6$ GeV, in the cases considered here. Recall again that the charm quarks do not come to a stop if only the collisional energy-loss is included [21].

Thus we conclude that the radiative energy-loss of heavy quarks produced initially in relativistic nuclear collisions plays a dominant role in pulling them to a stop in the QGP both at RHIC and LHC energies. Their final momentum distribution will then be determined by the temperature at which the hadroniza-

---

The fugacities are defined [18] as $f_i = \lambda_i \tilde{f}_i$, where $f$ is the distribution of the partons and $\tilde{f}$ is the corresponding equilibrium distribution.
tion takes place. This could be the temperature of the mixed phase, if such a phenomenon takes place.

We may add that Svetitsky and co-workers [4] have actively investigated such a scenario. In their work the large drag coefficient arises due to a large value of $\alpha_s \approx 0.6$ and a fully equilibrated plasma, even though only the collisional energy-loss is included.

Let us return to the discussion of the momentum distribution of charm quarks. (Similar considerations hold for bottom quarks.) It is expected that the momentum distribution of charm quarks will be reflected in the momentum distribution of charmed mesons, which are produced during the hadronization of the QGP and whose correlated decay will provide a back-ground to dileptons from quark annihilation. We see immediately that a look at the $p_\perp$ distribution of these leptons may help us to isolate the two contributions, as they should be very different for the two sources. Shuryak [23], Lin et al. [24], and Kömper et al. [25] have argued that the correlated charm decay back-ground for dileptons may be suppressed if the energy-loss of charm quarks is calculated charm decay back-ground for dileptons may be suppressed if the energy-loss of charm quarks is taken as $1-2 \text{ GeV}/\text{fm}$. Our study lends a strong support to their conclusion which were obtained by attributing an arbitrarily assumed value for the energy-loss.

In conclusion, we have estimated the radiative energy-loss of heavy quarks propagating in a quark-gluon plasma. This, along with the (fairly small) collisional energy-loss acts as a strong drag force on heavy quarks, which pulls them to a stop even in a chemically equilibrating and cooling plasma. This ensures that the momenta of the resulting charm mesons will be determined by the hadronization temperature. The correlated decay of such charm mesons will then no longer pose a back-ground for dileptons having their origin in the quark-antiquark annihilation at least at large invariant mass. This separation could even be made easier by measuring the $p_\perp$ distribution of the lepton pairs. In a future publication we shall report the result of the transverse hydrodynamic flow of the plasma on these conclusions [19].

We gratefully acknowledge helpful discussions with Bikash Sinha. One of us (D.K.S.) would like to thank for the hospitality of the University of Bielefeld, where part of this work was done.

References

[22] An estimate of radiative energy-loss of charm quarks was also given in Ref. [21]. The results given here were obtained after more elaborate checks on the applicability of the approximation used by Gunion and Bertsch [13] and putting $p^2 \approx 3T$ while getting Eq. (5). The results for radiative energy-loss plotted in Figs. (2) of Ref. [21] unfortunately are wrong by a factor of $1/k_c \sim 5$ by which they should be multiplied. The results involving the collisional energy-loss given there are correct.
A new phenomenon in heavy ion inelastic scattering: the towing mode

J.A. Scarpaci a, D. Beaumel a, Y. Blumenfeld a, Ph. Chomaz b, N. Frascaria a, J. Jongman a, D. Lacroix b, H. Laurent a, I. Lhenry a, F. Maréchal a, V. Pascalon-Rozier a, P. Roussel-Chomaz b, J.C. Roynette a, T. Suomijärvi a, A. Van der Woude c

a Institut de Physique Nucléaire, IN2P3-CNRS, 91406 Orsay, France
b GANIL, BP 5027, 14076 Caen Cedex 5, France
c Kernfysisch Versneller Instituut, 9747 AA Groningen, The Netherlands

Received 15 December 1997; revised 23 March 1998

Abstract

The inelastic scattering of 40Ar on 58Ni has been studied at 44 MeV per nucleon incident energy in coincidence with light particle emission. Besides the well known mechanisms of inelastic excitation, pick up break up and nucleon knock out, a new phenomenon has been observed, giving rise to fast forward moving particles with specific angular correlations. This newly observed mechanism seems to be a generic phenomenon present for various projectile-target combinations and incident energies. Its contribution to the inelastic spectrum has been extracted and a tentative interpretation is given.

PACS: 25.70.Bc; 25.70.Hi; 25.70.Pq; 24.10.-i

Inelastic scattering of heavy ions has been extensively used to study nuclear structure. In particular, the excitation and decay of giant resonances and multiphonon states has been investigated to infer information on their microscopic nature [1]. In addition to collective excitations of the target, the decay of which gives rise to isotropically emitted low energy particles, two other processes are known to contribute to the inclusive inelastic scattering spectrum: nucleon knock-out and pick-up break-up. At incident energies around 50 MeV per nucleon, these different reactions can be well separated by measuring light charged particles in coincidence with the inelastically scattered projectiles [2]. The quasi elastic knock out process gives rise to particles emitted around the direction of the recoiling target nucleus. This mechanism, known to be a major contribution to inelastic scattering of protons and α-particles, has been seen to persist in heavy ion induced scattering, but the shape and the strength of its contribution to the ejectile energy spectrum has not been quantita-
tively extracted so far. The pick-up break-up mechanism, described in Ref. [3], is a two step process. The projectile picks up a nucleon from the target, leaving it in a 1-hole configuration, and then breaks up by emitting one nucleon, thus feeding the inelastic channel. The velocity of the emitted nucleon is boosted by the projectile velocity, leading to the emission of the nucleon in a narrow cone around the direction of the scattered projectile, which is in our case detected at angles smaller than 5 degrees. The opening angle of this cone is approximately 30 degrees in the case of proton emission from a mass 40 projectile at 50 MeV per nucleon [2], and is expected to be smaller in the case of neutron emission due to the absence of the Coulomb barrier. In the ejectile energy spectrum, this process gives rise to an approximately flat shaped contribution over an excitation energy interval centered slightly above the bombarding energy per nucleon and with a width determined by the specific decay characteristics of the system [3].

During the course of systematic studies of heavy ion inelastic scattering [4–6], it was found that in addition to the three processes alluded to above, there is a fourth mechanism occurring as a result of the collision process. This new process gives rise to particles emitted in the forward hemisphere exhibiting very specific angular correlations. However, in these previous experiments, the angular coverage for light particles was not optimized for the study of this mechanism. To enhance our understanding of the inelastic channel in heavy ion reactions, a detailed study of this process is called for.

We report here on the 58Ni(40Ar, 40Ar + n or p) inelastic scattering reaction performed at 44 MeV per nucleon, in which the particle decay to the ground state (GS) and the low-lying excited states of the daughter nucleus exhibits an anomalous behaviour. We find that the most energetic decay particles are focused in a narrow cone in a direction which is not compatible with any of the known reaction mechanisms described above. In-plane and out-of-plane angular correlations between the 40Ar ejectile and the emitted particles will be presented in order to characterize the new mechanism, named “towing mode”. Its contribution to the inclusive inelastic spectrum will be extracted and an interpretation will be suggested.

The experiment was performed at the GANIL facility by bombarding a 1 mg/cm² 58Ni target with the 44 MeV per nucleon 40Ar beam. Protons were detected in 28 cesium iodide (CsI) elements of the multidetector array PACHA [7]. Most of these detectors were located on two rings centered on the target at a distance of 12 cm. The detectors were placed out of the horizontal plane at out of plane angles of −30° and −50°. Unambiguous identification of protons was obtained through shape analysis of the CsI pulse. The energy resolution was 2% above 20 MeV and around 6% for protons of energy less than 10 MeV.

Neutrons were detected with the liquid scintillator array EDEN [8] consisting of 45 modules positioned at about 1.75 meters from the target at angles running from 35° to 290°. In-plane angles are counted clockwise, 0° corresponding to the beam direction. The neutron identification was performed through a pulse shape analysis and the energy was obtained by a time of flight measurement that yielded an energy resolution of about 500 keV for 6 MeV neutrons and 2 MeV for 16 MeV neutrons. All the data in coincidence with neutrons were corrected for the detection efficiency according to Ref. [8].

The ejectiles were detected with the SPEG spectrometer [9] centered at Θlab = 3° and equipped with its standard detection system, consisting of two position sensitive drift chambers, an ionization chamber and a plastic scintillator. Both the horizontal and vertical opening angles are ±2°. An unambiguous identification of the ejectile was obtained from an energy loss measurement in the ionization chamber combined with a time of flight measurement which yielded the charge and the mass over atomic charge ratio of the ejectile, respectively. During the analysis a software gate was set on 40Ar to select the inelastic channel. Using the momentum and the scattering angle of the ejectile deduced from the reconstructed trajectory and applying two body kinematics, we calculate for each event an “apparent” excitation energy E’. This energy corresponds to the total deposited energy in the case of inelastic excitation of the target. The energy resolution obtained was 800 keV. The missing energy is calculated as follows: Emiss = E’ − Ep − Erec where Ep and Erec are the kinetic energies of the detected particle and of the recoiling nucleus respectively. In
the case where only one light particle is emitted in the reaction, the missing energy is related to the final state energy of the target nucleus $E_{fs}$ by $E_{fs} = E_{miss} + Q$, where $Q$ is the reaction $Q$-value.

Fig. 1 shows the missing energy spectra obtained for proton (a and c) and neutron (b and d) decay respectively, for excitation energies above 30 MeV and for two groups of detectors, one at backward angles in the laboratory frame and the other one in the forward direction, on the same side of the beam as the ejectile as sketched on the figure. For both detector groups, a large bump around 30 MeV is observed which corresponds to the decay of the target. At forward angles (between 30° and 85°) an important additional contribution feeding the GS and the first excited states of the daughter nuclei is present.

In order to study the angular correlations of decay particles populating the low-lying states of these nuclei, a 4 MeV wide gate was set on missing energies around the GS of the daughter nuclei, 12 MeV and 9 MeV for $^{57}$Ni and $^{57}$Co respectively, as shown on Fig. 1. These correlations are presented on Fig. 2 for the full acceptance of the spectrometer. The neutron angles are shown in the laboratory system. Most of the proton detectors are located out-of-plane and we have chosen to plot the proton angular distributions as a function of the angle in the horizontal plane. The correlations are plotted for apparent excitation energies between 12 MeV (16 MeV for protons) and 80 MeV. Besides an almost isotropic component which could result from the decay of the target, a very strong peak is observed around +40° on the same side of the beam as the ejectile. A fit, including a flat background and a Gaussian, has been performed. The amplitude of the Gaussian peak is much larger for neutrons than for protons. It is located at angles larger than what is expected for pick up break up reactions and on the opposite side of the beam from knocked out nucleons. In the proton angular correlation we also observe, located around −60° on the opposite side of the beam from the ejectile, the recoil protons coming from the elastic scattering from the hydrogen contained in the target, and possibly also from a contribution of the knock-out process.

Further insight into the emission mechanism of the particles can be given by azimuthal angular correlations. Such correlations could be extracted as the ejectile was detected at azimuthal angles located between $\Phi_{eject} = \pm 50°$, where $\Phi_{spher} = 0$ is taken in the horizontal plane on the right side of the beam where the ejectile is detected. Most proton detectors were positioned outside of the horizontal plane, $\Phi_{prot}$ around −37° for detectors on the right side of the beam and $\Phi_{prot}$ around −143° for detectors on the left side of the beam. Fig. 3(a) presents the spherical angles of ejectiles, $\Phi_{eject}$ and $\Theta_{eject}$, in

![Fig. 1. Missing energy spectra for excitation energies above 30 MeV for forward (a) and (b) and backward (c) and (d) emitted particles. The missing energy spectra in coincidence with neutrons were corrected for the neutron detection efficiency. For each particle type, the two spectra are normalized to the same detector solid angle. The right side of the figure shows a sketch of the ejectile velocity after scattering, the emitted particle velocity and the velocity of the undetected recoiling target. Dashed vertical lines represent the gates as set for the angular correlations presented in Fig. 2.](image-url)
coincidence with backward emitted protons coming from the target decay. The shape of this scatterplot simply reflects the square entrance of the SPEG spectrometer located on the right side of the beam. Even though the protons were detected out of the horizontal plane, the azimuthal angular distribution is symmetric around $\phi^\text{spher}_{\text{eject}} = 0$ demonstrating that there is no correlation between the particles arising from the target decay and the direction of the scattered projectile. Fig. 3(b) shows the same two-dimensional spectrum for ejectiles in coincidence with protons stemming from the elastic scattering on the hydrogen contaminant in the target. As the proton is detected around $\theta^\text{spher}_{\text{prot}} = 56^\circ$ and $\phi^\text{spher}_{\text{prot}} = -143^\circ$ (on the left side of the beam), the ejectile is observed around $\phi^\text{spher}_{\text{eject}} = +37^\circ$ as expected for a binary reaction. The insert of Fig. 3(b) schematically depicts the velocity of the elastically scattered ejectile (full dot) and of the emitted proton (open dot) in the plane perpendicular to the beam velocity (crossed circle). Fig. 3(c) shows the scatterplot of ejectiles in coincidence with protons feeding the GS and the first excited states of

![Fig. 2. Neutron and proton angular correlations for particles feeding the GS of the daughter nucleus between 7 and 11 MeV in the missing energy spectrum for the protons and 10 and 14 MeV for the neutrons (see Fig. 1)) and for apparent excitation energies below 80 MeV. The dashed lines are the result of a fit with a Gaussian peak plus a constant background. The characteristics of the Gaussian are reported.](image)

![Fig. 3. Angular distribution ($\phi^\text{spher}_{\text{eject}}$ and $\theta^\text{spher}_{\text{eject}}$) for the ejectile in coincidence with backward emitted protons (a), protons coming for the scattering on the hydrogen of the target (b) and protons feeding the GS of $^{57}\text{Co}$ and emitted around $\theta^\text{spher}_{\text{prot}} = 40^\circ$ and $\phi^\text{spher}_{\text{prot}} = -43^\circ$ (c). Inserts of panels (b) and (c) schematically depict the velocities of the ejectile (full dot) and of the proton (open dot) in the plane perpendicular to the beam axis (represented as a crossed circle), for the elastic scattering (b) and for the new mechanism (c).](image)
the daughter nucleus ($^{57}$Co) and emitted around $\Theta_{\text{prot}} = 45^\circ$ and $\Phi_{\text{prot}} = -43^\circ$, corresponding to the mechanism of interest. Contrary to Fig. 3(a), a strong correlation between the proton and the ejectile azimuthal angles is observed. However here, the ejectile and the proton are both detected below the horizontal plane, at about the same azimuthal angle, $\Phi_{\text{eject}} = -43^\circ$, as schematically summarized in the insert.

These strong correlations, both in-plane and out-of-plane, between the particle and the ejectile velocities, combined with the information that the particle energies are on the average close to half of the beam energy per nucleon, suggest an interpretation of the particle emission mechanism in terms of an aborted pick-up. It seems that the particle is extracted from the target and towed along for a short while by the ejectile. Therefore we name this mechanism “Towing Mode”.

The contribution of this mechanism to the inclusive inelastic spectrum can be extracted by comparing the inelastic spectra in coincidence with forward emitted particles (between $+35^\circ$ and $+70^\circ$), where the contribution from towing mode particles is present, with the one in coincidence with backward emitted particles (between $+120^\circ$ and $+170^\circ$), which stem solely from target decay, with no cut on the ejectile angle. The extracted contribution, shown on Fig. 4, is peaked at an apparent excitation energy of about 30 MeV for both neutrons and protons and extends as far as 80 MeV. It is interesting to note that the neutron contribution is an order of magnitude larger than the proton contribution. This could be due to the structure of $^{58}$Ni composed of two valence neutrons around the closed shell $^{56}$Ni core or to the Coulomb barrier.

A tentative normalization to the inclusive inelastic spectrum has been performed as follows. The angular correlations were extracted with no constraint on the excitation energy nor on the missing energy. They were then fitted by a Gaussian plus a constant background, and the Gaussian curves were integrated assuming an azimuthal angle running between $\pm 50^\circ$ which corresponds to the azimuthal opening angle of the spectrometer, giving the total number of counts for the mechanism (this is based on the assumption that the projectile and the emitted particle have the same azimuthal angle). Fig. 4 shows the inclusive inelastic spectrum (top) as well as the extracted contribution for the new mechanism, for both neutron (solid line) and proton (dashed line) emission. The shape of the spectrum reflects the kinetic energy distribution of the light emitted particles. Also shown is the inelastic spectrum in coincidence with backward emission of neutrons or protons normalized to $4\pi$ that reflects the excitation of the target (open dot spectrum). Above 40 MeV the contribution of the new mechanism is of the same order of magnitude as that of target excitation. In the reaction studied, it contributes for about 20% of the total inelastic cross section in the excitation energy region of 20 to 100 MeV. The additional strength observed in the inclusive spectrum must be mainly due to the pick-up break-up mechanism [10].

This experiment shows that in this new process the target nucleus is left in a 1-hole state (cf. missing energy spectra, Fig. 1), just as in the pick-up break-up and the knock-out processes. However the transferred nucleon does not form a short-lived unbound system with the projectile and thus is not fully boosted to the projectile velocity. The angular correlations (Fig. 2 and Fig. 3) show that the direction of the particle is strongly correlated to that of the ejectile but does not correspond to what is expected either for the pick-up break-up process or for the
knock-out process. The “Towing Mode”, seems to be a general phenomenon in heavy ion scattering since a similar component of fast forward moving particles has been observed in many other inelastic scattering experiments, \(^{208}\text{Pb} (^{17}\text{O}, ^{17}\text{O} + \text{n}) [5], ^{40}\text{Ca} (^{8}\text{O}, ^{10}\text{Ca} + \text{p}) [4], ^{48}\text{Ca} (^{20}\text{Ne}, ^{20}\text{Ne} + \text{n}) [11], ^{90,94}\text{Zr} (^{36}\text{Ar}, ^{36}\text{Ar} + \text{n}) [6] at incident energies between 40 and 84 MeV per nucleon. Even at incident energies as high as 400 MeV per nucleon, fast forward moving neutrons were observed in the reaction \(^{84}\text{Xe} (^{17}\text{O}, ^{17}\text{O} + \text{n}) [12]. The authors of Ref. [12] interpreted this fast component as knock-out of neutrons in the far-out region of the nucleus, i.e. in the nuclear stratosphere. However, the out-of-plane correlation measured in our experiment tells us that the fast particles that we observe are not coming from a knock-out process since it would then resemble the elastic scattering azimuthal correlation, hence completely opposite to what we observe.

To investigate if a model can qualitatively explain our observations, a calculation has been performed to infer the evolution of a particle wave function when two potentials brush past each other. The time dependent Schrödinger equation was solved in a two dimensional space using Wood-Saxon potentials of radii \(r_s = 1.2 \text{ A}^{1/3}\) and diffuseness \(a_s = 0.5\) Fermi. It describes the evolution of a wave function initially in the target potential at rest when the projectile passes by at a given impact parameter. Since we are dealing with very peripheral reactions in which both the target and the projectile are very little disturbed (they remain mostly cold after the collision), we assumed constant wells. Fig. 5 shows the probability density for a 2s wave function initially bound in the target potential by 8.7 MeV, when the projectile has passed by at an impact parameter of 8 Fermi at an incident energy of 44 MeV per nucleon. We observe that there is a large probability that the particle remains in the target (73%), and some probability that it is transfered to the projectile (1.4%). But most interesting is the sizeable probability (21%) of particle emission at an angle around 50° on the same side of the beam as the projectile, qualitatively reproducing the experimental observation. This calculation will be more extensively discussed in a forthcoming paper.

Summarizing, we have observed and characterized a new phenomenon in the inelastic channel of heavy ion collisions at intermediate energies. This process consists in the emission of a fast nucleon leaving the residual nucleus in the GS or a low-lying hole state. The angular characteristics of the emitted particle, which are inconsistent with what is expected from either the pick-up break-up or the nucleon knock-out mechanisms, support an interpretation in terms of an aborted transfer where the nucleon in towed by the projectile for a very short period of time and then emitted into the continuum. A calculation solving the time dependent Schrödinger equation predicts a similar emission of particles in the same spatial region. Future experiments, with a more complete angular coverage for light particles, should investigate the dependence of this phenomenon on the projectile-target combination, the incident velocity and the initial angular momentum of the emitted particle in order to further our quantitative understanding of this new phenomenon.

References

Precise measurements of the neutron magnetic form factor

H. Anklin a, L.J. deBever a, K.I. Blomqvist c, W.U. Boeglin c, R. Böhm c, M. Distler c, R. Edelhoff c, J. Friedrich c, D. Fritschi a, R. Geiges a, J. Götz a, A. Honegger a, P. Jennewein c, J. Jourdan a, M. Kahrau c, M. Korn c, H. Kramer c, K.W. Krygier c, G. Kubon a, V. Kunde c, A. Liesenfeld c, G. Masson a, K. Merle c, R. Neuhausen c, E.A.J.M. Offermann c, Th. Petitjean a, Th. Pospischil c, M. Potokar b, L.M. Qin a, A.W Richter c, A. Rokavec c, G. Rosner c, P. Sauer c, S. Schardt c, I. Sick a, S. Sirca c, Ph. Trueb a,2, M. Tuccillo a, B. Vodenik b, A. Wagner c, Th. Walcher c, G. Warren a, S. Wolf c, J. Zhao a, M. Zeier a, B. Zihlmann a

a Dept. für Physik und Astronomie, Universität Basel, Basel, Switzerland
b Institute Jožef Stefan, University of Ljubljana, Ljubljana, Slovenia
c Institut für Kernphysik, Johannes Gutenberg-Universität, Mainz, Germany

Received 3 December 1997; revised 27 February 1998
Editor: J.P. Schiffer

Abstract

The neutron magnetic form factor $G_{m n}$ has been determined via a measurement of the ratio of cross sections $D(e,e'n)$ and $D(e,e'p)$. The absolute detection efficiency of the neutron detector was measured with high accuracy using tagged neutrons produced from $H(n,p)n$ elastic scattering by means of a high intensity neutron beam. This approach minimizes the model dependence and improves upon the weakest points of previous experiments. Data in the range $q^2 = 0.2-0.8 \text{ (GeV/c)}^2$ with uncertainties of $<2\%$ are presented. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

Precise data on the dependence of the nucleon form factors on the four momentum transfer $q^2$ provide detailed information on the inner structure of the nucleon and serve as a sensitive test for models of the nucleon. The proton form factors are known with excellent precision over a large range of momentum transfers. Due to the lack of a free neutron target, data for the neutron are of much poorer quality. This is true for both the electric form factor $G_{e n}$ and to a somewhat lesser degree for the magnetic form factor $G_{m n}$.  

370-2693/98/$19.00 \copyright 1998$ Elsevier Science B.V. All rights reserved.

Pll: S0370-2693(98)00442-0
\( G_{mn} \) has been determined mostly from quasi-elastic \( D(e,e') \) cross sections [1–9]. The extraction of \( G_{mn} \) requires a longitudinal/transverse separation and a subtraction of the (dominant) proton magnetic contribution. The uncertainties resulting from the deuteron model, exchange currents (MEC), and final state interactions (FSI) are greatly amplified by the two subsequent subtractions. Mostly due to these theoretical uncertainties, the accuracy of \( G_{mn} \) is no better than \( \sim 20\% \). An alternative approach, quasi-coincidence \( D(e,e'p) \) [1,2,5], relies upon the absence of a recoiling proton \( \vec{p} \) as identification of \( (e,e'n) \) and requires an excellent understanding of the reaction mechanism, which is not at hand. More recently \( G_{mn} \) was extracted from a measurement of the asymmetry in \( ^3\text{He}(e,e') \) [10] and from a cross section measurement of \( D(e,e'n) \) [11]. The asymmetry experiment makes use of a new technique and the \( D(e,e'n) \) experiment avoids the subtraction of the proton contribution. However, both measurements suffer from the limitations of past, the sensitivity to the structure of the nucleus employed as a neutron target.

To minimize the sensitivity to the nuclear structure one can extract \( G_{mn} \) from the ratio \( R = (d\sigma(e,e'n)/d\Omega)/(d\sigma(e,e'p)/d\Omega) \) on the deuteron in quasi-free kinematics [12–15]. The \( e-p \) cross section is well known and the ratio is insensitive to the wave function. Corrections due to FSI and MEC are small and well understood when restricting the acceptance to the quasi-elastic peak. The difficulty is shifted to the experiment: the precise determination of the absolute efficiency \( \eta \) of the neutron detector.

In other experiments the neutron detection efficiency has been determined employing the \( D(\gamma,p)n \), \( H(\gamma,\pi^+)n \), and \( H(e,\pi^+)e'n \) reactions to produce a ‘beam’ of tagged neutrons [12–14,11]. However, it is not practical to determine \( \eta(x,y,T_n) \) as a function of the neutron energy \( T_n \) and the location of impact on the detector \( x,y \) because of the limitations in the intensity of these ‘beams’. In addition, for the case of \( H(e,\pi^+)e'n \) used in [14], the 3-body final state introduced serious ambiguities [16]. Much of the complications in determining \( G_{mn} \) originates in the difficulties of such efficiency measurements. In fact, it has been argued that the recent data [14] suffer from a systematic error in the efficiency determination [16].

In our approach we employ a method [15] which improves upon the weakest point of the experiments: the determination of the neutron efficiency. Precise measurements of \( \eta(x,y,T_n) \) and a detailed study of the detector response are possible when using high intensity neutron beams available at the proton-beam facilities. This has been demonstrated in a pilot experiment [15] where \( G_{mn} \) was determined with an accuracy of 1.7%.

2. Measurements at PSI

In the present work, the determination of the neutron detection efficiency was performed using a beam of 100–500 MeV neutrons available at the Paul Scherrer Institut (PSI) [17]. Neutrons are produced via the \( \text{C(p,n)} \) reaction with a neutron flux (integrated over the entire energy range) of \( 10^8 \) s for a 10 \( \mu \)A proton beam and a 9 mm beam collimator. As in [15] the tagged high intensity neutron beam was produced via the \( H(n,p)n \) reaction, scattering the neutrons from a 1 cm thick liquid \( H_2 \)-target. The recoil protons were detected with \( \Delta E \) and \( E \) plastic scintillators recording the amplitude and the time-of-flight (TOF). Four multi-wire proportional chambers (MWPC) determined the target coordinates with an accuracy of \( \pm 1.5 \) mm and the recoil angles with an accuracy of \( \pm 0.2^\circ \).

The energy spectrum of the incident neutron beam was very different from the one used in [15] which had important consequences. The advantage of a wide energy spectrum was a simultaneous efficiency measurement over a large range of energies. However, due to the RF-structure of the proton beam (50 MHz) and the flight distance of 24 m, the determination of the neutron energy via TOF relative to the RF was not unique. This ambiguity was removed by measuring as well the TOF from the \( H_2 \)-target to the \( E \) plastic bars. The precise calibration of the TOF was achieved with \( \gamma \)-rays produced in the neutron production target, converted to charged particles in a thin Pb-target at the position of the \( H_2 \)-target, and detected with the \( \Delta E \) scintillator.

The knowledge of the incident and the recoil energy, as well as the recoil angles provided a redundant determination of the energy and position of the tagged neutron beam which was free from
contributions of background reactions. The accuracy of the tagged energy was 2.1–15.6 MeV for the energy range (120–460 MeV) of the tagged neutrons and the position resolution on the detector surface was 7.0 mm. As in [15] the hadron detector was placed in this neutron beam and its efficiency measured. The detector consisted of two 10 cm thick plastic converters, \( E_x \) and \( E_y \), preceded by 3 thin \( \Delta E \) counters to identify the incident nucleon. Except for the opening towards the target, the detector was shielded with 10 cm thick lead walls. Lead absorbers of 1–20 mm thickness were placed at the entrance window of the detector. These absorbers are needed for the determination of \( H(n,p) \) and the \( D(e,e'n) \) measurements in order to absorb low energy photons. Use of the identical configuration at PSI ensured that absorption effects of neutrons were automatically included in the efficiency determined.

When using a neutron beam with a continuous energy spectrum, tagged scattered neutrons of the full range in energy are produced at all scattering angles. The high intensity of the tagged neutron flux allowed to study \( \eta(x,y,T_n) \) as a function of \( x,y,T_n \) simultaneously, without moving the detector as was done in [15]. The determination of the absolute efficiency distribution \( \eta(x,y,T_n) \) is necessary because the detector illumination is unavoidably different for the \( H(n,p)n \) and the \( D(e,e'n) \) data taking. An example of \( \eta(x,y,T_n) \) is shown in Fig. 1.

The dominant source of the variation of \( \eta(x,y,T_n) \) due to the light collection efficiency was measured separately using \( H(n,p) \) recoil protons, as it needs to be known with good spatial resolution near the detector edges. This was done by placing the detector on the recoil arm and replacing the \( E \) scintillators by a second thin \( \Delta E \)-detector to provide a tag for the recoil protons which does not depend on the neutron detector. The spatial resolution is provided by the four MWPC. The knowledge of \( \eta(x,y,T_n) \) allows to match the efficiency \( \eta \) needed to determine \( R \) from the yield ratio with minimal corrections (see Table 1). If the efficiency \( \eta \) is measured only for a central illumination of the detector at the center value \( T_n \) of the \( (e,e'n) \)-measurement this correction can be as large as 14%.

Tagged recoil protons were also used for several additional corrections which are needed in the determination of \( R \) over the entire energy range. In particular, multiple scattering and absorption effects of the protons, which are the dominant correction of \( R \), were measured for all Pb-absorbers employed with an accuracy of 0.7% (Table 1). Whereas at low energy multiple scattering mostly reduced the \( D(e,e'p) \) yield, absorption effects dominated in the present experiment due to the thicker lead absorbers and the higher proton energies employed. Table 1 summarizes all the measured corrections with its uncertainties.

3. Measurements at MAMI

The measurements of the \( D(e,e'n) \) and the \( D(e,e'p) \) yield in quasi-elastic kinematics for \( q^2 = 0.235, 0.504, 0.652, \) and 0.784 (GeV/c)^2 were performed at the Mainz Microtron (MAMI) [18]. As in [15] the \( (e,e'n) \) and the \( (e,e'p) \) yields were measured simultaneously which made the ratio independent of the luminosity, dead time effects, and the efficiency of the electron arm. The essential difference to [15] is the 100% duty factor of MAMI which is needed in order to achieve an acceptable signal-to-noise ratio for the \( (e,e'n) \) measurement at higher \( q^2 \). In the present experiment the signal-to-noise ratio for neutrons was better than 150, even at the highest \( q^2 \).

The measurements were performed with a 0.1–1.2 \( \mu A \) cw electron beam incident on a cylindrical 2 cm thick liquid deuterium target cell with 6 \( \mu m \) HAVAR windows. The incident beam energy was
555 MeV for the lowest \( q^2 \)-point and 855 MeV for all higher \( q^2 \)-measurements. Spectrometer A [19] with a solid angle limited to 21 msr detected the scattered electron in coincidence with the recoiling hadron. The hadron detector covered an angular range of \( \pm 78 \) mrad horizontal and vertical resulting in a solid angle of 24.3 msr. The detector covered 50–90% of the Fermi cone depending on kinematics.

Considerable care was taken to obtain identical cuts on the energy deposited in the detectors. \( H(e,e'p) \) at MAMI and \( H(n,p) \) at PSI were used to obtain absolute calibrations of the amplitude spectrum. In addition, each converter was monitored for both gain variations and baseline shifts using two temperature compensated LED’s whose light output was in turn monitored by very stable PIN-diodes [15]. The energy scale of the converters was identical to an accuracy of 0.6% in the PSI and the MAMI runs.

The off-line analysis of the \((e,e'n)\) and the \((e,e'p)\) yields were very similar to the analysis described in [15]. The threshold used in the definition for neutrons in \( E_f \) was set at 25 MeV. The number of counted neutrons was corrected for the efficiency of the veto condition and for misidentified protons due to inefficiencies of the veto counters. Depending on the kinematics the determined veto efficiency corrections are 1.6–10.8%. The correction for misidentified protons ranged from 0.8% to 4.9%. The three independent measurements for the number of neutrons determined with the three different combinations of \( \Delta E \) counters agreed within 0.3%. The uncertainty in these corrections are small relative to the statistical error of the number of neutrons and are included in the error of the yield ratio as shown in Table 1.

Protons were counted in the TOF spectrum of \( E_f \) with a coincident time signal in at least one of the three \( \Delta E \) counters. The number of protons was corrected for the inefficiency of the \( \Delta E \) counter used. The final number of protons, obtained using different \( \Delta E, E_f \) combinations, agreed within 0.08%.

### Table 1

<table>
<thead>
<tr>
<th>( q^2 )(GeV/c(^2))</th>
<th>0.235</th>
<th>0.504</th>
<th>0.652</th>
<th>0.784</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error of the yield ratio</td>
<td>St</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
</tr>
<tr>
<td>Illumination matching</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
</tr>
<tr>
<td>Threshold calibration</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
</tr>
<tr>
<td>Proton losses</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
</tr>
<tr>
<td>(p,n)-correction</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
</tr>
<tr>
<td>Error of ( \eta )</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
</tr>
<tr>
<td>( T_n ) and ( x,y ) uncertainty</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
<td>sy</td>
</tr>
<tr>
<td>( R )</td>
<td>0.234</td>
<td>0.366</td>
<td>0.415</td>
<td>0.445</td>
</tr>
<tr>
<td>Relative error of ( R ) in %</td>
<td>( \pm 1.5 )</td>
<td>( \pm 2.0 )</td>
<td>( \pm 2.0 )</td>
<td>( \pm 2.3 )</td>
</tr>
</tbody>
</table>

### 4. Evaluation of \( R \)

From the measured ratio of detected neutrons and protons, the ratio \( R \) of the \( D(e,e'n) \) and \( D(e,e'p) \) cross sections was determined, making use of the carefully determined absolute efficiency of the detector. Two efficiency measurements bracketed the measurement of the yield ratio at MAMI in order to check the reliability of \( \eta \). Consistent results for \( \eta \) are found, and the resulting values of \( R \) are independent of the applied \( E_f \) threshold. Table 1 summarizes errors and applied corrections, all of which are based on measurements, and the final results for \( R \).

To extract \( G_{nn} \) from the measured \( R \) values we have to go beyond the plane wave impulse approximation (PWIA) and take into account the effects of FSI, MEC, and isobar currents (IC). The results for \( R^{\text{theo}} \) of Arenhövel [20], obtained with the Paris potential, were used. The n–p final state interaction, mostly charge exchange scattering, gives the largest contribution of 1.3–4.0% to \( R \) depending on kinematics. The contributions from MEC and IC are of...
Table 2
Results for $G_{mn}$ relative to the dipole form factor $G_D = (1 + q^2/0.710)^{-2}$. The error on $G_{mn}$ includes both the experimental contribution (table 1, $\sigma_{e-n}$, $G_{en}$) and the ones due to theory $\text{en}$. $\sigma_{e-n}$ and $G_{en}/(\mu_n G_D)$. $q^2 (\text{GeV}/c)^2$ $0.235$ $0.504$ $0.652$ $0.784$

| $[(R_{\text{PWIA}}/R)^{\text{theo}} - 1](\%)$ | $-4.9 \pm 0.6$ | $-2.8 \pm 0.4$ | $-2.2 \pm 0.3$ | $-1.9 \pm 0.2$ |
| Contribution of $G_{se}(\%)$ | $-2.0 \pm 0.6$ | $-2.4 \pm 0.7$ | $-1.8 \pm 0.5$ | $-1.2 \pm 0.4$ |
| $\sigma_{e-p}/\sigma_D$ | $0.938 \pm 0.009$ | $0.958 \pm 0.009$ | $0.987 \pm 0.008$ | $1.022 \pm 0.006$ |
| $G_{en}/(\mu_n G_D)$ | $0.962 \pm 0.009$ | $1.032 \pm 0.012$ | $1.037 \pm 0.012$ | $1.043 \pm 0.012$ |

order 0.5% and relativistic effects are negligible. The total corrections are listed in Table 2 as a deviation $D = (R_{\text{PWIA}}/R)^{\text{theo}} - 1$ from the PWIA-value. The dependence of $D$ on the nucleon–nucleon potential was studied at $q^2 = 0.235$ and $0.504 (\text{GeV}/c)^2$ where the correction is largest. The variation found when using the Bonn-R-space, the Argonne V14, and the Nijmegen potential is well within the systematic uncertainties (FSI $\pm 15\%$, MEC $\pm 40\%$, IC $\pm 60\%$) used to calculate the errors of Table 2.

The resulting value $R_{\text{PWIA}} = (\sigma_{e-n}/\sigma_{e-p})_{\text{PWIA}} = R \cdot (1 + D)$ is the experimental ratio of the $e-n$ cross section (which is essentially given by $G_{mn}$) relative to the $e-p$ cross section corrected for non-PWIA contributions. The contribution of the neutron electric form factor to $R$ is small; it introduces only a small additional uncertainty despite the poor knowledge of $G_{en}$ (see Table 2).

To evaluate the $e-p$ cross section $\sigma_{e-p}$, we used the world’s supply of $\sigma_{e-p}$-data in a range of $0.5 \text{ fm}^{-1}$ around the desired $q$. In this range of $q$, we used a parameterization with a relative $q$-dependence as given by the Höhler fit [21] to $G_{ep}$ and $G_{mp}$, with the overall normalizations fitted to the

![Fig. 2](image-url)

Fig. 2. Top figure shows the available data on $G_{mn}$ relative to the Dipole parameterization. Shown are the data of [15] (solid squares), Gao et al. [10] (open squares), Markowitz et al. [11] (open diamonds), Bruins et al. [14] (open circles), and the present work (solid circles). Data measured before 1990 are only shown by the error bar. The bottom figure shows the data of Anklin et al. [15] (solid squares), and of the present work (solid circles), in comparison to various model calculations. Solid: Mergell et al. [22], dot: Meissner [23], dash-dotted: Eich [24], dash: Schlumpf [25], dash-dot-dotted: Lu et al. $r_{\text{Bag}} = 0.9 \text{ fm}$ [26].
world data. The statistical errors of the data have been treated in the standard way, while the systematic errors have been accounted for by changing the data of each set by its error, refitting, and adding the changes due to systematic errors in quadrature. The resulting $e - p$ cross sections are listed in Table 2 together with the final results for $G_{mn}$, which are plotted in Fig. 2. Both the proton and neutron form factors differ from the crude empirical expression $G_q = (1 + q^2 / 0.710)^{-2}$ used to remove the dominant $q^2$-dependence in Table 2 and Fig. 2. The discrepancy of the present results of $G_{mn}$, with data from other recent experiments [14,11] has already been discussed elsewhere [15,16].

The data of the present work and of Anklin et al. [15] show significant differences to both the non-relativistic constituent quark model calculation by Eich [24], and the relativistic version by Schlumpf [25]. Similar differences are observed when comparing the data to a recent cloudy bag model calculation by Lu et al. [26]. The data are also compared to the minimal vector dominance model by Meissner [23] and the recent calculation by Mergell et al. [22] based on a fit of the proton data using dispersion theoretical arguments. None of the calculations describe the data satisfactorily.

5. Conclusion

In the present experiment we determined $G_{mn}$ in the $q^2$-range from 0.23 (GeV/c)^2 to 0.78 (GeV/c)^2. We achieved a substantial decrease of the uncertainties. The improvement in the knowledge of $G_{mn}$ is due to the fact that we use an approach that depends the least on the input of theory; this becomes possible by performing the needed calibrations using a high-intensity tagged neutron beam, and by using an electron beam with 100% duty factor. The results, listed in Table 2, indicate that accuracies of $< 2\%$ of $G_{mn}$ can be reached. This increase of accuracy allows for a detailed analysis of the electromagnetic structure in terms of theoretical model calculations.

References

Cosmic strings in low mass Higgs cosmology

Warren B. Perkins \(^{a,1}\), Anne-Christine Davis \(^{b,2}\)

\(^a\) Department of Physics, University of Wales Swansea, Singleton Park, Swansea SA2 8PP, Wales, UK
\(^b\) Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 9EW, UK

Received 2 March 1998; revised 24 March 1998
Editor: P.V. Landshoff

Abstract

A class of grand unified theories with symmetry breaking scale of order \(10^{16}\) GeV have a Higgs particle with mass in the TeV scale. The cosmology of such theories is very different from usual. We study the cosmic strings obtained in such theories. These strings are much fatter than usual and their mass per unit length is reduced, resulting in a significant reduction in their cosmological effects. We also study the temperature evolution of such models. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

In recent years there has been much interest in unified fields theories, in particular, those involving supersymmetry. In a wide class of supersymmetric grand unified theories some of the phase transitions occur in the TeV range, despite the gauge symmetry breaking scale being of order \(10^{16}\) GeV. This occurs in a rather generic class of supersymmetric theories in which the scalar field potential has ‘flat directions’. Such theories can arise in superstring theories \(^1\). The associated scalar fields also have mass in the TeV range, usually arising from very small Higgs self-couplings. However, the gauge bosons have mass of order the grand unified scale.

Supersymmetry is looking increasingly likely to be involved in some underlying theory that unifies the interactions, whether it is a superstring theory or a supersymmetric grand unified theory. In such theories, the so-called flat directions are also a common feature, arising in a class of theories where the gauge symmetry is broken with a so-called F-term. Such models could solve the cosmological moduli problem. Whilst the physical content of such theories is understood, it is only recently that the cosmology has started to be investigated \(^2\)–\(^4\).

In a very innovative paper \(^2\) Lyth and Stewart considered the cosmology of models with low mass Higgs particles. They showed that the evolution of the universe was radically different from usual. In addition to the usual inflationary period their model had a period of thermal inflation followed by a ‘cold big bang’ at TeV scales. This was succeeded by Higgs particle decay and a ‘hot big bang’ at around \(10\) MeV just prior to nucleosynthesis.

In this paper we investigate the cosmology of such theories, considerably extending previous work. We consider topological defects formed in such the-
ories, in particular cosmic strings. In these theories the string profiles are considerably modified by the relatively small Higgs mass. We display the string profiles in Section 2 for both the Higgs and gauge fields and calculate the string width and mass per unit length. We show that there is a modification factor depending on the logarithm of the ratio of the scalar and gauge particle masses. This modification results in the strings being ‘fat’, with decreased mass per unit length. In Section 3 we discuss string dynamics, showing that they are produced after the period of friction domination, but otherwise evolve as normal strings. The resulting cosmology of such strings is also considered. We show that, if the gauge symmetry breaking scale is around $10^{16}$ GeV, then the strings are less cosmologically significant than usual. We also discuss microphysical affects of such strings. In Section 4 we comment on the temperature-time relation in our model and compare it with the temperature evolution discussed previously. By considering the decay of the Higgs particles in a half-life model, we show that the dramatic results of [2] are modified considerably.

We summarize our results in Section 5.

2. The structure of low Higgs mass strings

In order to consider cosmic strings [6] in grand unified theories with low mass Higgs particles we can model the relevant features of the Higgs potential after supersymmetry breaking by the usual Mexican hat with a very small Higgs self coupling. We thus consider an effective Abelian Higgs model with Lagrangian,

$$\mathcal{L} = (D_\mu \phi)^\dagger D_\mu \phi - \frac{1}{4} F_{\mu \nu} F^{\mu \nu} - \Lambda (\phi^* \phi - \eta^2/2)^2.$$

It should be emphasized that this potential is only appropriate for tree level calculations. Loop corrections within the Abelian Higgs model would usually necessitate a degree of fine tuning to protect the very small Higgs self coupling we are considering. However, in the context of the underlying supersymmetric theory this fine tuning is not a problem. We can consistently work with the Abelian Higgs model at the tree level with whatever couplings we choose.

We construct a Nielsen-Olesen vortex solution in the usual manner by setting

$$\phi = \phi(r) e^{i\theta}$$

and considering only $A_\theta$ non-vanishing. The energy per unit length of string is then given by

$$\epsilon = 2\pi \int_0^\infty r dr \left[ \phi_r^2 + \phi_\theta^2 \left( \frac{1}{r} - e A_\theta \right)^2 + \frac{1}{2} \left( A_{\theta,\theta} + A_\theta/r \right)^2 + \lambda (\phi^2 - \eta^2/2)^2 \right].$$

It is convenient to introduce the scaled variables:

$$\phi = \frac{\eta}{\sqrt{2}} \tilde{\phi}, \quad A_\theta = \eta \tilde{A}, \quad r = \frac{x}{\eta e}.$$

In terms of these variables the energy per unit length reduces to,

$$\epsilon = \pi \eta^2 \int_0^\infty x dx \left[ \tilde{\phi}_r^2 + \tilde{\phi}_\theta^2 \left( \frac{1}{x} - \tilde{A} \right)^2 + (\tilde{A}_{,r} + \tilde{A}/x)^2 + B (\tilde{\phi}^2 - 1)^2 \right],$$

where $B = \lambda/2 e$.

In order to obtain a model for the string we take the following forms for the string profiles:

$$\tilde{\phi} = \begin{cases} \beta x^\gamma, & 0 \leq x \leq \beta^{-1/\gamma} \\ 1, & x > \beta^{-1/\gamma} \end{cases}$$

$$\tilde{A} = \begin{cases} \frac{x}{X^2}, & x \leq X \\ 1, & x > X \end{cases}$$

These approximations allow the gauge and scalar cores to have different radii and allow the scalar field to either delay its approach to its asymptotic form (large $\gamma$) or move to it more rapidly (small $\gamma$).

With these forms for the profile functions the energy per unit length becomes,

$$\epsilon = \pi \eta^2 \left[ \frac{\gamma}{2} + \beta^2 X^{2\gamma} \left( \frac{1}{2\gamma} - \frac{1}{\gamma + 1} + \frac{1}{2\gamma + 4} \right) \right. + \frac{2}{X^2} + B \beta^{-2/\gamma} \left( \frac{1}{2} - \frac{1}{\gamma + 1} + \frac{1}{4\gamma + 2} \right).$$

...
We now wish to vary the parameters so as to minimize the string energy. We first gain a rough idea of how the parameters vary with $B$ and then proceed to a more careful minimization.

For small values of $B$, that is small Higgs self coupling, we expect the energy to be small. This requires $\gamma$ to be small, $X$ to be large and $\beta^2/\gamma$ to be small in order to suppress the individual terms in the energy. Assuming that each term in the energy decreases as the inverse of some scale $l$ as $B$ becomes small, we have

$$\gamma \sim \frac{1}{l}, \quad X \sim \sqrt{l}, \quad \beta \sim \frac{1}{l}$$

and

$$B \sim l^{-1+2l}$$

As an approximation we consider

$$l = \frac{\log(1/B)}{2\log(1/B)}.$$ 

With this value of $l$, $B\beta^{-2/\gamma}$ indeed decreases more rapidly than $1/l$, thus our model for the string profile functions gives an energy per unit length that decreases at least as rapidly as

$$2\log\log(1/B) \quad \frac{\log(1/B)}{\log(1/B)}.$$

We need to be slightly more careful about the minimization in order to determine how the core sizes scale with $B$. We introduce some shorthand,

$$O = \frac{1}{2\gamma} - \frac{1}{\gamma+1} + \frac{1}{2\gamma+4},$$

$$P = \frac{1}{2} - \frac{1}{\gamma+1} + \frac{1}{4\gamma+2},$$

then vary the energy with respect to each of the parameters. Varying first with respect to $X$ we find

$$X = \left( \frac{2}{\beta^2\gamma O} \right)^{1/2+2\gamma}.$$ 

Eliminating $X$ and varying with respect to $\beta$, we find that the value of $\beta$ is given by

$$\beta^2 = \left[ \frac{BP}{2} \left( \frac{Q\gamma}{2} \right)^{-\frac{1}{1+\gamma}} \right]^{\frac{1}{1+2\gamma}}.$$ 

We saw above that $\gamma$ is small for small values of $B$, so we expand our expression for the energy about $\gamma = 0$ and keep only the leading terms;

$$\epsilon = \pi\eta^2 \left[ \frac{\gamma}{2} + \frac{B\gamma}{2\gamma} \right].$$

Minimizing with respect to $\gamma$ yields the constraint,

$$\frac{K^2}{\log^2 B} + (K-1)e^K = 0,$$

where $K = \gamma \log B$. There is a positive solution, $K = 1 + O(1/\log^2 B)$, but this gives an unphysical negative value to $\gamma$. For large $\log^2 B$, the negative solution is

$$K = -\log(\log^2 B).$$

Finally we have

$$\gamma = \frac{\log(\log^2(1/B))}{\log(1/B)} = \frac{2\log\log(1/B)}{\log(1/B)}.$$

While this coincides with the form for $\gamma$ we obtained by the naive argument, the forms of $X$ and $\beta$ are different. Working back through the constraints we find

$$\beta = B\frac{\gamma}{\log(1/B)}, \quad X = 2\log(1/B).$$

Fig. 1 shows a comparison of the actual Higgs and gauge field profiles of a low Higgs mass string with the model profiles for $B = 10^{-4}$. Thus, the mass per unit length for these type of cosmic strings is

$$\mu = \frac{\eta^2}{\log(B^{-1})}.$$ 

Hence, for symmetry breaking scale $\eta$ the mass per unit length is reduced by the logarithmic factor. A
similar result was obtained numerically in ref [3], though the full profile functions were not obtained. For small $\lambda$ this reduction factor can be an order of magnitude, hence affecting the cosmological predictions in this model. Similarly, the gauge and scalar cores are vastly different. The scalar core size is set by the the inverse Higgs' mass, whilst the gauge core is increased by $\log B^{-1}$. Hence, in these low mass Higgs models the string width is considerably fatter than in the usual case. That these strings are much fatter was recognised in ref [4], though the reduction in $\mu$ was not realised.

3. The evolution and cosmology of low Higgs mass strings

In the standard string model there are two distinct periods of string evolution, depending on the signifi-
We note that the plasma length scale above which friction is dominant is set by [5]

\[ l_f = \frac{\mu}{\sigma \rho} \]

where \( \mu \) is the mass per unit length of string, \( \rho \) is the energy density of scatterers and \( \sigma \) is the scattering cross-section per unit length. While the mass per unit length of these strings is suppressed by a logarithmic factor, the scattering cross-section is unchanged. The dominant process is Aharonov-Bohm scattering, with the cross-section being determined by the momentum of the scattering particles, \( \sigma \propto T^{-1} \).

In standard string models friction domination ends well before the electroweak transition. The analysis is similar in this modified picture, with a logarithmic correction appearing due to the lower mass per unit length. The main difference is that strings are formed much later, at around the time of the electroweak transition. Thus strings form well after the period of possible friction domination. Otherwise their evolution is simply that of normal string with a suitably reduced mass per unit length.

Whilst the usual Aharonov-Bohm scattering cross-section is the same for our fat strings as for the normal GUT strings the same is not true for the inelastic cross-section. For example, the baryon violating catalysis cross-section will depend on the string radius in realistic GUT models, unlike the case in toy \( U(1) \) models. This dependence arises from the group generators in the string gauge field and it also depends on the scattering particle. The full details are shown in [7]. Consequently, the catalysis cross-section will be bigger for fat strings than the corresponding usual GUT strings. This could result in some erasing of a primordial baryon asymmetry.

The cosmological effects of cosmic strings are determined by the parameter \( G\mu \). The anisotropy in the microwave background radiation, the density fluctuations and the gravitational lensing are all proportional to \( G\mu \). With the logarithmic reduction factor, as given in the previous section, the cosmological effects of these fat strings will be just those of ordinary strings formed at a slightly lower energy scale. This reduction in the gravitational effects of the strings prevents them from playing a major role in structure formation, unless the unification temperature is slightly raised by a corresponding amount. Similarly, the power spectrum is likely to be the same shape as for the usual GUT scale strings. These strings are likely to evade constraints arising from microwave background and large scale structure measurements. As these strings are appearing in an inflationary model, this the lack of string density perturbations is not a problem, perturbations arising during the inflationary era can seed large scale structure formation. There has been recent interest in mixed models where the density perturbations are produced by both cosmic strings and inflationary perturbations. If the cosmic strings are of the fat type discussed here then their contribution to density perturbations will be reduced unless unification is raised to compensate for the logarithmic factor.

The decay of cosmic string loops can result in the production of cosmic rays and also produce a baryon asymmetry. In [8,9] it was shown that the decay of GUT scale string loops could account for the resulting baryon asymmetry, particularly in models with a low freeze out temperature. The corresponding case for our fat strings is rather more complicated. This is because the analysis in [8] was performed at the Ginsberg temperature where thermal fluctuations can no longer restore the symmetry. However, since the fat strings under consideration are produced in first order phase transitions, it seems more appropriate to consider the transition temperature itself. In which case, we have a massive enhancement factor coming from the increase in the string width. The results of ref [8] are enhanced by a factor of \( \lambda^{-1} \), in which case this model can easily account for the observed baryon asymmetry.

Similarly, there will be a large amount of cosmic rays produced by the decaying string loops, mainly in the form of TeV Higgs particles. There may also be cosmic rays produced by the infinite string network [10]. Since the particles emitted by the network are going to be mainly TeV scale Higgs particles and their decay products [11], they will evade the bounds of ref [10].

However, since the dominant microphysical effects occur shortly after string formation, these will be diluted by any subsequent reheating. If there is sufficient reheating both the baryon asymmetry produced and the cosmic ray flux will be diluted, and
could be diluted below observational limits. We discuss this in the next section.

4. The temperature-time relationship for late transitions

The extremely low Higgs mass in these models delays the breaking of the GUT symmetry until the temperature is of order the Higgs mass, \( \sim 1 \) TeV. Even with the very flat Higgs potential corresponding to such low self couplings, the vacuum energy density is of order \( 10^{36} \text{ GeV}^4 \) and is much greater than the radiation energy density (of order \( 10^{12} \text{ GeV}^4 \) at symmetry breaking). Following symmetry breaking the vacuum energy density is converted into Higgs particles and the universe enters a phase of matter domination [2]. In [2] it is argued that this phase lasts for the lifetime of a typical Higgs particle, of order \( 10^{23} \) GeV, then the Higgs particles decay, leading to a reheating of the Universe. This reheating produces a standard radiation dominated epoch that encompasses the final period of nucleosynthesis.

While this approach provides a reliable estimate of the temperature at which radiation domination recommences, the temperature evolution during the Higgs dominated phase is rather different. As discussed in [12], it is more appropriate to think of the Higgs lifetime as a half-life and consider a continuous transfer of energy from the matter to radiation fields.

If \( r_d \) is the probability of a Higgs particle decaying per unit time, the evolution of the radiation and matter energy densities are given by

\[
\frac{d}{dt} \rho_r = -4H \rho_r + \rho_m r_d, \quad \frac{d}{dt} \rho_m = -3H \rho_m - \rho_m r_d.
\]

These expressions are valid for both a second order transition [13] and for the first order transition we expect at weak coupling. Bubble collisions produce a sea of ‘soft’ quanta [14], with energy less than the corresponding instantaneous reheat temperature. Now, even instantaneous conversion of the Higgs’ potential energy to radiation would only give a reheat temperature of around \( 10^9 \) GeV, much less than the mass of the particles mediating Higgs decay. Thus the Higgs decay rate will be that of static Higgs particles.

Following [12], we can understand how the energy densities evolve by considering some approximate solutions. In a flat, matter dominated Universe, the evolution of the matter energy density is given by,

\[
\frac{d}{dt} \rho_m = -3\sqrt{\frac{2}{\pi}} G \rho_m \rho_m - \rho_m r_d. 
\]

and we have \( \rho_m = x e^{-r_d t} \) with

\[
x_f^{-1/2} = x_i^{-1/2} - \frac{3}{r_d} \sqrt{\frac{8}{3}} \pi G [e^{-r_d t_i/2} - e^{-r_d t/2}].
\]

To leading order in \( r_d t \) this gives

\[
\rho_m = [x_i^{-1/2} + \frac{1}{2} \sqrt{\frac{8}{3}} \pi G (t_f - t_i)]^{-2}.
\]

Assuming that \( x_i \) is very large for \( t_i \) close to zero, we have the standard form,

\[
\rho_m = \left[ \frac{1}{\frac{\sqrt{8}}{\pi} G t} \right]^{-2}.
\]

As expected, for \( r, t \ll 1 \) the decay of the Higgs particles has little effect on their energy density. However, significant energy is transferred into radiation. With \( \rho_m \) varying as \( t^{-2} \), we can solve for the radiation energy density,

\[
\rho_r = 4 \frac{r_d 1}{15 \frac{8}{3} \pi G t}.
\]

The particular integral varies as \( t^{-1} \) and so quickly dominates the complementary function. Physically, the initial radiation is diluted and redshifted, quickly leaving only the radiation from the decaying Higgs particles, thus the temperature decreases as \( t^{-1/4} \) once the initial transient has decayed [12]. A possibility not discussed in [12] is that this transient can allow the temperature to increase at the onset of Higgs decay, leading to a prompt reheating. The amount of prompt reheating depends on the Higgs lifetime and the ratio of scalar to radiation energy densities immediately after the symmetry breaking. If the radiation energy density is very low and the Higgs decay is rapid, there will be an immediate injection of energy into a cold Universe, leading to prompt reheating.
Assuming that these small $r_d t$ forms persist until matter-radiation equality, we have equality at $r_d t = 5/3$. Now, the Hubble parameter is given by $H = 2/3t$, thus at this time the Higgs' half-life and the expansion time-scale are comparable and Higgs decay becomes important in determining $\rho_m$. Following equality the remaining Higgs particles decay rapidly, but the energy liberated is only of order the radiation energy density and there is little further reheating.

We can compare the entropy generated in this approximation with that obtained by assuming that all of the Higgs' decay at $t = 1/r_d$. In both cases we have a phase of matter domination with $R \propto t^{2/3}$.

Fig. 2. The Higgs (solid line) and radiation energy densities as functions of time.

Fig. 3. The temperature as a function of time for continuous Higgs decay (solid line) and instantaneous decay.
which ends at $t \sim 1/t_d$. This is followed by a phase of radiation domination with an initial radiation energy density,

$$\rho_r \sim \frac{r_d^2}{3 \pi G}.$$

The entropy generation and temperature at equality are similar in both approximations, but the temperature evolutions are very different as is the amount of entropy produced subsequent to any given intermediate temperature.

The forms of $r_s$ and $r_m$ for $r_d = 10^{-22}$ GeV and the initial conditions given above are shown in Fig. 2. It can be seen that there is a small amount of prompt reheating, but then the universe cools monotonically with matter domination ending just before $\rho_r$ drops below about $10^{-12}$ GeV$^{-4}$ and nucleosynthesis begins.

The temperature evolutions for continuous and instantaneous Higgs decay are compared in Fig. 3. The most significant differences between the two arise for temperatures above $T_{\text{reheat}}$, these temperatures are only attained once and occur before much of the entropy generation in the instantaneous model. In the continuous decay model these temperatures are reached later and, particularly in the case of temperatures close to $T_{\text{reheat}}$, there is much less entropy generation at subsequent epochs.

In the instantaneous picture, the temperature increases by a factor of around $10^{10}$ at reheating. Thus the ratio of the relic density to the entropy density, $n/s$, for any relic produced at temperatures above the reheat temperature is reduced by a factor of around $10^{30}$ at reheating. In our continuous case, the scale factor has the usual matter dominated form, $R \propto t^{2/3}$, while the temperature has the somewhat unusual form, $T \propto t^{-1/4}$. The relic density to entropy density ratio then has the form,

$$\frac{n}{s} \propto \frac{R^{-3}}{T^3} \propto t^{-1.25}.$$

For a relic formed at a temperature $T_x$ at time $t_x$, we have a dilution factor, d.f., of

$$\text{d.f.} = \left( \frac{t_{\text{reheat}}}{t_x} \right)^{1.25} = \left( \frac{T_x}{T_{\text{reheat}}} \right)^{5/4}.$$

For processes occurring at around the TeV scale, the dilution factor is around $10^{30}$ as in the instantaneous case. However, if the process occurs at a lower temperature, the dilution factor is reduced.

In the baryogenesis scenario of [8], the baryon asymmetry was sufficient to account for that required by nucleosynthesis. Our mechanism discussed in the previous section enhances this by the factor $\lambda^{-1}$. However, the dilution factor we have found is considerably greater than this. Consequently, the resulting baryon asymmetry will be diluted below that required by nucleosynthesis. A very late baryogenesis mechanism is needed in this class of models in order to avoid dilution. Similarly, emitted cosmic rays are likely to be significantly diluted in this class of models.

5. Conclusions

We have considered the properties of cosmic strings in theories where the Higgs self-coupling is small. Such theories arise in supersymmetric, grand unified theories with flat directions. We have shown the resulting strings are much fatter than usual, with vastly different gauge and scalar field core sizes. The gauge core is increased by a logarithmic factor, whilst the scalar core is the inverse Higgs mass. Similarly, the mass per unit length of the string is reduced by the logarithmic suppression factor relative to usual cosmic string models. This alters the cosmology of such strings. The gravitational properties are reduced by this factor. Consequently, if such strings are to play a role in large scale structure then the unification temperature needs to be increased by a corresponding amount. This makes such models less attractive from a particle physics viewpoint, though they could still arise from superstring theories. Similarly, the increase in string width changes their microphysical properties. It results in an increased baryon catalysis cross-section and a vast increase in the number of particles produced by loop decay. Whilst this latter effect could produce the observed baryon asymmetry and result in massive production of cosmic rays from such strings, it is likely that such observable effects will be diluted by the subsequent reheating.

The evolution of the universe is greatly modified in this class of theories, with the universe undergoing
a late reheating. Using a realistic model for the decay of the TeV scale Higgs particle, we have shown how the evolution is modified. There is prompt reheating following the phase transition, the universe then undergoes a period of matter domination as it cools monotonically until radiation domination begins just before nucleosynthesis. This scenario is rather different from that proposed in [2] where all Higgs particles were taken to decay at the typical lifetime of order $10^{23}$ GeV$^{-1}$. In particular, relics produced at temperatures above the reheat temperature suffer significantly less dilution if the Higgs decay is continuous than they would if the decay was instantaneous.

Acknowledgements

This work was supported in part by PPARC. This work is in part the result of a scientific network supported by the European Science Foundation. W.B.P. would like to thank R.A.W. Gregory for useful discussions.

References

String theory formulation of anti-de Sitter black holes

Danny Birmingham

Department of Mathematical Physics, University College Dublin, Belfield, Dublin 4, Ireland

Received 11 February 1998
Editor: P.V. Landshoff

Abstract

It is shown that the five-dimensional anti-de Sitter black hole is a supersymmetric solution of the low-energy field equations of type IIB string theory compactified on an Einstein space. A statistical interpretation of the mass dependence of the entropy can be obtained from considerations of the three-dimensional BTZ black hole. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

The study of black holes in string theory has been the source of recent developments in our understanding of the origin of the Bekenstein-Hawking entropy formula [1–3]. In many cases, an important role is played by supersymmetry, and the observation that D-branes carry charge for the RR fields of string theory [4]. In three dimensions, a particularly interesting example of a black hole, known as the BTZ black hole, has been constructed [5,6], see [7] for a review. The construction is based upon the observation that by performing a quotient of three-dimensional anti-de Sitter space (adS), one obtains a spacetime with the properties of a black hole. It was observed recently [8] that the microscopic entropy can be understood for certain black holes whose near-horizon geometry is locally equivalent to adS. This observation is based on the fact that adS has an asymptotic symmetry algebra consisting of left and right Virasoro algebras [9,10]. In particular, the entropy of the extreme BTZ black hole, viewed as a supersymmetric solution of heterotic string theory without the use of RR fields, was computed in [11].

In [12], a higher-dimensional generalization of the BTZ construction was provided. The essential idea is to take a quotient of anti-de Sitter space, yielding a black hole with topology $\mathbb{R}^{d-1} \times S^1$, and an explicit construction in five dimensions was presented. An interesting aspect of these higher-dimensional black holes is that the horizon is a circle, and thus one might hope to be able to understand their entropy from a lower-dimensional point of view. We show that the five-dimensional black hole is a supersymmetric solution of type IIB string theory compactified to $\text{adS}_d \times K_5$, where $K_5$ is a compact internal Einstein space, with only the field strength for the RR 4-form excited. At extremality, the black hole has a non-zero horizon length, proportional to the square root of the mass. A statistical interpretation of this mass dependence can be achieved by relating it to the entropy of the three-dimensional BTZ black hole.

© 1998 Elsevier Science B.V. All rights reserved.

PII: S0370-2693(98)00425-0
2. Compactification of type IIB string theory on \text{adS}_5

The five-dimensional anti-de Sitter black hole is obtained as a quotient of \text{adS}_5 \times K_5, to which it is locally equivalent [12]. To obtain this five-dimensional black hole as a supersymmetric solution of type IIB string theory, we seek a compactification to \text{adS}_5 = \text{K}_5,5, where \text{K}_5 is a compact internal space. The black hole is then obtained by performing the necessary identifications. We write the 10-dimensional coordinates as \( x^M = (x^\mu, y^m) \), with \( \mu = 0,1,2,3,4, \) and \( m = 5,6,7,8,9 \), and we follow the conventions of [13].

The covariant field equations for type IIB string theory have been determined in [14]. As observed in [14], the desired compactification can be obtained by setting all the bosonic fields to zero, except for the metric and anti-self-dual 5-form field strength. In this case, the only non-trivial equation of motion is given by

\[
R_{MN} = \frac{1}{4^2 \times 6} F_{MN} N_1 N_2 N_3 N_4 F_{N_1 N_2 N_3 N_4} ,
\]

where the 5-form is taken to be anti-self-dual. Thus, we have

\[
F_{M_1 M_2 M_3 M_4 M_5} = - \frac{1}{5!} \epsilon_{M_1 M_2 M_3 M_4 M_5} F_{N_1 N_2 N_3 N_4 N_5} ,
\]

with \( \epsilon^{0123456789} = 1/\sqrt{-g} \). We wish to obtain a product metric of the form

\[
g_{\mu\nu} = g_{\mu\nu}(x), \quad g_{mn} = g_{mn}(y), \quad g_{\mu m} = 0 .
\]

The required solution follows by choosing the ansatz [14]

\[
F_{\mu_1 \mu_2 \mu_3 \nu_1 \nu_2 \nu_3} = Q^2 \epsilon_{\mu_1 \mu_2 \mu_3 \nu_1 \nu_2 \nu_3} ,
\]

\[
F_{m_1 m_2 m_3 m_4 m_5} = - Q^2 \epsilon_{m_1 m_2 m_3 m_4 m_5} .
\]

It is then straightforward to see that

\[
R_{\mu \nu} = - \frac{Q^2}{4} g_{\mu \nu} , \quad R_{mn} = \frac{Q^2}{4} g_{mn} ,
\]

with \( R_{\mu m} = 0 \), and we note that the 10-dimensional Ricci scalar vanishes.

In order to establish the supersymmetry of the solution, we must show that the Killing spinor equations are satisfied. Namely, we must show that the supersymmetry variations of the fermionic fields vanish in the compactified background. The relevant terms in the supersymmetry transformations take the form [13,14]

\[
\delta \psi_M = \nabla_M \epsilon + \frac{i}{4 \times 480} \Gamma_{M_1 M_2 M_3 M_4} \epsilon F_{M_1 M_2 M_3 M_4} ,
\]

\[
+ \frac{i}{4 \pi} \left( \Gamma_{M_1 M_2 M_3} - 9 \delta_{M_1} \Gamma_{M_2 M_3} \right) \epsilon^* F_{M_1 M_2 M_3} ,
\]

\[
\delta \lambda = i \Gamma_M \epsilon^* \left( \frac{\nabla_M \phi}{1 - \phi^* \phi} \right) - \frac{i}{24} \Gamma_{M_1 M_2 M_3} \epsilon F_{M_1 M_2 M_3} .
\]

A representation of the Dirac matrices which is relevant to the \( 5 + 5 \) split is given in [15,16]. The 10-dimensional Dirac matrices are denoted by \( \Gamma^A \) and satisfy

\[
\{ \Gamma^A, \Gamma^B \} = 2 \eta^{AB} ,
\]

with signature \( \eta^{AB} = (- + \cdots +) \). The representation is given by

\[
\Gamma^A = ( \Gamma^a, \Gamma^\alpha ) = ( \gamma^a \otimes 1_4, \gamma^\alpha \otimes \Sigma^a ) ,
\]

where the spacetime matrices are 8-dimensional, and the internal matrices are 4-dimensional, satisfying

\[
\{ \gamma^a, \gamma^\beta \} = 2 \eta^{a\beta} , \quad \{ \Sigma^a, \Sigma^\beta \} = 2 \delta^{ab} ,
\]

with signature \( \eta^{a\beta} = (- + + + +) \). The \( \gamma^6 \) matrix is given explicitly by

\[
\gamma^6 = \begin{pmatrix} 1_4 & 0 \\ 0 & -1_4 \end{pmatrix} ,
\]

and satisfies

\[
\{ \gamma^6, \gamma^a \} = 0 , \quad (\gamma^6)^2 = 1_8 .
\]

The spacetime Dirac matrices can be written as [15,16]

\[
\gamma^0 = \begin{pmatrix} 0 & i \sigma^1 \otimes 1_2 \\ i \sigma^1 \otimes 1_2 & 0 \end{pmatrix} ,
\]

\[
\gamma^1 = \begin{pmatrix} 0 & \sigma^3 \otimes 1_2 \\ \sigma^3 \otimes 1_2 & 0 \end{pmatrix} ,
\]
\[ \gamma^2 = \begin{pmatrix} 0 & \sigma^2 \otimes \tau^2 \\ \sigma^2 \otimes \tau^2 & 0 \end{pmatrix}, \]
\[ \gamma^3 = \begin{pmatrix} 0 & \sigma^2 \otimes \tau^1 \\ \sigma^2 \otimes \tau^1 & 0 \end{pmatrix}, \]
\[ \gamma^4 = \begin{pmatrix} 0 & \sigma^2 \otimes \tau^3 \\ \sigma^2 \otimes \tau^3 & 0 \end{pmatrix}. \] (13)

where \( \sigma^i \) and \( \tau^i \) are the Pauli matrices. The internal matrices are
\[ \Sigma^5 = \sigma^1 \otimes 1_2, \quad \Sigma^6 = \sigma^3 \otimes 1_2, \quad \Sigma^7 = \sigma^2 \otimes \tau^2, \]
\[ \Sigma^8 = \sigma^2 \otimes \tau^1, \quad \Sigma^9 = \sigma^2 \otimes \tau^3, \] (14)
with
\[ \Gamma^{ab} = \gamma^{a\beta} \otimes 1_4, \quad \Gamma^{ab} = 1_8 \otimes \Sigma^{ab}. \] (15)

A matrix which also enters the analysis is defined by
\[ J = \begin{pmatrix} 0 & 1_4 \\ 1_4 & 0 \end{pmatrix}. \] (16)

and satisfies \([J, \gamma^a] = 0\).

We write the 10-dimensional spinor as \( \epsilon(x, y) = \eta(x) \otimes \chi(y) \). The Killing spinor equations then become
\[ \delta \psi_\mu = 0 \Rightarrow \nabla_\mu \eta = -\frac{Q}{8} J_{\mu} \eta, \]
\[ \delta \psi_m = 0 \Rightarrow \nabla_m \chi = -i \frac{Q}{8} \Sigma_m \chi, \] (17)

with the constraint \( iJ \gamma^\beta \eta = \eta \). The integrability conditions implied by (17) are then precisely the conditions (5). Thus, we have obtained \( \text{ads}_5 \times K_5 \) as a supersymmetric solution of the low-energy equations of motion of type IIB string theory. The amount of supersymmetry present in five dimensions is then determined by the holonomy of the internal space \( K_5 \). One should also note that the constraint on \( \eta \) is consistent with the chirality on \( \epsilon \).

Defining the 10-dimensional chirality operator as
\[ \Gamma^{11} = \Gamma^0 \Gamma^1 \cdots \Gamma^9, \] (18)

with \((\Gamma^{11})^2 = 1\), we find that
\[ \Gamma^{11} = iJ \gamma^6 \otimes 1_4. \] (19)

Hence,
\[ \Gamma^{11} \epsilon = \epsilon \Rightarrow iJ \gamma^6 \eta = \eta. \] (20)

3. The five-dimensional anti-de Sitter black hole

Recently, a higher-dimensional generalization of the BTZ black hole was obtained [12]. The construction is analogous to the three-dimensional case, whereby a certain quotient of anti-de Sitter space is constructed with the properties of a black hole. In particular, the five dimensional case was explicitly constructed. The important point for our purposes is that the topology of the black hole is \( \mathbb{R}^{d-1} \times S^1 \), with the horizon being given by the \( S^1 \) factor. This is to be contrasted with the topology \( \mathbb{R}^2 \times S^{d-2} \) of a Schwarzschild black hole. The \( \text{ads}_5 \) black hole is parametrized by two parameters, its mass and angular momentum, and as shown in [12], these can be defined by relating the construction to a Chern-Simons supergravity theory for the supergroup \( SU(2|N) \) [17].

The line element of the \( \text{ads}_5 \) black hole can be written in the form [12]
\[ ds^2 = -\left(\frac{r^2 - r_+^2}{\ell^2}\right) \frac{\ell^2}{r_+^2} \cos^2 \theta dt^2 + \left(\frac{r^2 - r_+^2}{\ell^2}\right)^{-1} dr^2 + r^2 \frac{\ell^2}{r_+^2} d\phi^2 + \left(\frac{r^2 - r_+^2}{\ell^2}\right)^{-1} \left(d\theta^2 + \sin^2 \theta \, d\chi^2\right). \] (21)
in the range \(-\infty < t < \infty, r_+ < r < \infty, 0 < \theta < \pi, 0 < \chi < 2\pi\). The location of the horizon is specified by \( r = r_+ \). We note that at the horizon the angular \( (\theta, \chi) \) part of the line element also vanishes. In order to introduce angular momentum, one makes the replacements
\[ t \rightarrow \frac{r_+}{\ell^2} t - r_+ \phi, \quad \phi \rightarrow \frac{r_+}{\ell^2} t - \frac{r_+}{\ell} \phi, \] (22)
and identifies points along the new angular coordinate \( \phi \sim \phi + 2\pi n \), with \( r_+ < r_+. \) The line element then becomes
\[ ds^2 = dt^2 \left[ \frac{r^2 \ell^2}{r_+^2} - \left(\frac{r^2 - r_+^2}{\ell^2}\right) \cos^2 \theta \right] + \left(\frac{r^2 - r_+^2}{\ell^2}\right)^{-1} dr^2 \]
In order to relate the mass and entropy of the adS₅ black hole to the BTZ case, we need a relationship between the three- and five-dimensional Newton constants. Within string theory, it is natural to identify $k = \alpha'/8G₅$, where $G₅$ is the five-dimensional Newton constant with dimensions $+3$ in length. The entropy is then given by

$$S_{BH} = (2\pi r_+ \alpha') / 4G₅. \tag{28}$$

However, in order of magnitude, we have $G₃ \sim G₅/\alpha'$. We thus see that the mass of the five-dimensional black hole is of the order of the mass of the BTZ black hole. Upon this identification, one then finds that the entropy of the adS₅ black hole is of the order of the BTZ entropy (which has a statistical interpretation [8,11,18]). In order words, this identification yields the correct mass dependence of the five-dimensional entropy. In this sense, the analysis is similar in spirit to the correspondence principle presented in [19,20]. However, further work is required in order to fix the numerical coefficient.

4. Conclusions

The construction of conserved charges for the five-dimensional anti-de Sitter black hole relied on the formulation of a Chern-Simons supergravity theory for the supergroup $SU(2|2\times N)$ [17]. In this regard, it is worth remarking that a computation of the entropy for the BTZ black hole from the point of view of three-dimensional Chern-Simons theory with boundary was provided in [18]. The computation is based essentially on the observation that the horizon dynamics on the boundary is controlled by a WZW model. It would be interesting to see if a similar computation in the $SU(2|2\times N)$ Chern-Simons theory sheds further light on the entropy. Indeed, there has been recent progress in understanding the connection between singleton field theories in anti-de Sitter space and the D-brane picture of black hole entropy [21]. Further considerations are contained in [22–24]. One should also mention that the self-dual 3-brane of type IIB string theory has been discussed in the context of vacuum interpolation between 10-dimensional Minkowski space and adS₅ $\times$ $S^5$ [25,26].
Another interesting observation is the fact that the Chern-Simons supergravity theory for the anti-de Sitter group can be constructed up to a maximal dimension of seven \cite{17}. One might therefore expect to be able to construct conserved charges for the seven-dimensional anti-de Sitter black hole, as in the five-dimensional case. Furthermore, the $\text{adS}_7$ black hole could then be obtained as a compactification $\text{adS}_7 \times K_4$ of 11-dimensional supergravity, with the 4-form field strength being proportional to the volume form of the compact internal space $K_4$ \cite{27,28}. In this way, the $\text{adS}_7$ black hole could be given an interpretation as a solution of the low-energy limit of M Theory, with the 2-brane field excited.

Acknowledgements

I would like to thank K. Sfetsos for valuable correspondence on the entropy discussion. This work was supported by Forbairt grant SC/96/603.

References

\cite{3} See, for example, G. Horowitz, Quantum States of Black Holes, gr-qc/9704072; J. Maldacena, Black Holes in String Theory, hep-th/9607235.
\cite{7} S. Carlip, Class. Quant. Grav. 12 (1995) 2853, gr-qc/9506079.
\cite{8} A. Strominger, Black Hole Entropy from Near-Horizon Microstates, hep-th/9712251.
\cite{12} M. Bañados, Constant Curvature Black Holes, gr-qc/9703040.
\cite{19} L. Susskind, Some Speculations about Black Hole Entropy in String Theory, hep-th/9309145.
\cite{22} J. Maldacena, The Large $N$ Limit of Superconformal Field Theories and Supergravity, hep-th/9711200.
\cite{23} S. Ferrara, C. Fronsdal, Conformal Maxwell Theory as a Singleton Field Theory on $\text{adS}_5$, IIB Three Branes and Duality, hep-th/9712239.
\cite{24} H.J. Boonstra, B. Peeters, K. Skenderis, Branes and anti-de Sitter spacetimes, hep-th/9801076.
Exact T-duality between calorons and Taub-NUT spaces

Thomas C. Kraan 1, Pierre van Baal 2
Instituut-Lorentz for Theoretical Physics, University of Leiden, P.O. Box 9506, NL-2300 RA Leiden, The Netherlands

Received 11 February 1998
Editor: P.V. Landshoff

Abstract

We determine all $SU(2)$ caloron solutions with topological charge one and arbitrary Polyakov loop at spatial infinity (with trace $2\cos(2\pi \alpha)\omega$), using the Nahm duality transformation and ADHM. By explicit computations we show that the moduli space is given by a product of the base manifold $R^1 \times S^1$ and a Taub-NUT space with mass $M = 1/\sqrt{8\alpha(1 - 2\alpha)}$, for $\omega \in [0, \frac{1}{2}]$, in units where $S^1 = R/Z$. Implications for finite temperature field theory and string duality between Kaluza-Klein and H-monopoles are briefly discussed. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

Properties of self-dual solutions to the Yang-Mills equations of motion have played an important role in understanding both the physical and mathematical properties of gauge theories. The last couple of years these solutions also feature prominently in the description of dualities in supersymmetric theories and in string theories, in particular for extensions to D-branes and M-theory.

We will present the calorons [1], which are instantons at finite temperature defined on $R^1 \times S^1$, in an explicit and simple form for topological charge one. The reader interested in the physical applications, like for finite temperature field theory, should skip the mathematically oriented introduction below and go directly to Section 2. Sections 3 and 4 can be skipped as well. A more detailed description will be published elsewhere.

We were inspired to pursue the case of calorons by a question posed one year ago by J. Gauntlett concerning the moduli space of calorons with non-trivial asymptotic behaviour of the Polyakov loop (non-trivial holonomy) [2]. These calorons appear as the non-trivial component of H-monopoles [3]. Oddly enough explicit solutions with non-trivial holonomy were not known. With trivial holonomy they can be obtained as an infinite periodic array of instantons, all oriented parallel in group space [1]. Recent work on the T-duality between Kaluza-Klein and H-monopoles in string theory [4] made us aware that our construction explicitly provides the classical duality transformation. It can be formulated without their embedding in string theory. We nevertheless hope this result can contribute to resolving some of the puzzles that seem to be involved in the relevant string dualities. There will be many experts better equipped than we are in addressing these stringy issues.

The Nahm transformation [5,6], also known as Mukai transformation [7] when considered as a map-
ping between holomorphic vector bundles, maps self-dual fields on $R^4/\Lambda$ to self-dual fields on $R^4/\Lambda'$. Here $\Lambda$ is an integer lattice and $\Lambda'$ is its dual. For the gauge group $U(N)$ this Nahm transformation interchanges the rank $N$ and the topological charge (also mapping the first Chern class to its Hodge dual), as follows from a family index theorem [8]. The family parameter is defined in terms of the moduli space of flat $U(1)$ connections, $\omega_2 = 2\pi iz_\mu dx_\mu$, which when added to the self-dual $U(1)$ connection does not change the curvature. This gives rise to a family of zero-modes for the chiral Dirac operator (the Weyl operator). The vector bundle defined over the (dual) space of flat connections thus obtained, has itself a self-dual connection. Monopoles, calorons, and instantons on $R^4$ can all be considered to arise from suitably chosen limits of lattices $\Lambda$. In particular for $R^4$ seen as a torus, all of whose sides are sent to infinity, the dual space is a single point (periods $L$ are mapped to $1/L$ and it is in this sense that the Nahm transformation is in fact a T-duality mapping [9]). This explains the algebraic nature of the Atiyah-Drinfeld-Hitchin-Manin (ADHM) construction [10] and can be used as most elegant and straightforward derivation [6,11,12]. It can be shown that the Nahm transformation is an involution; applied twice it gives the identity operation. Furthermore it preserves the metric and hyperKähler structure of the moduli spaces [8].

In the process of sending certain periods to infinity, boundary terms arise that destroy the self-duality of the Nahm bundle, but this can be repaired by suitably extending the Weyl operator on the dual space [6,11,13]. A particularly interesting feature arises on non-compact four dimensional manifolds for which infinity has the topology of $T^d$, where $d$ can be either 1, 2, or 3. These correspond respectively to instantons on $R^3 \times S^1$, $R^2 \times T^2$ and $T^3 \times R$. Note that in the latter case infinity actually contains two disconnected three dimensional tori [13]. As the solutions have finite action the connection at infinity is flat, parametrised by the Polyakov loops winding around the $d$ circles. Combined with the flat $U(1)$ connection that is added to perform the Nahm transformation, the Weyl operator reduced to the asymptotic $T^d$ generically has a gap. This is guaranteed to be the case as long as the combined flat $U(N)$ connection on $T^d$ is without flat factors [12], i.e. does not reduce to $U(1) \oplus U(N-1)$ with $U(1)$ trivial. For $SU(2)$ it is easily seen that the Weyl operator reduced to the asymptotic $T^d$ will have a zero eigenvalue at $2^d$ values of $x$. If the holonomy in the direction $i$ is in the center of the gauge group, the values for the components $x_i$ of these points will coincide. As soon as (at least one of) the Polyakov loops is non-trivial, the symmetry is spontaneously broken to $U(1)$. The zero-modes of the reduced Weyl operator lead to non-exponential decay of the zero-modes for the full Weyl operator and a partial integration, required in computing the curvature of the Nahm bundle, will pick up a boundary term. These boundary terms can only occur at the $2^d$ points mentioned above, and therefore are distributions, indeed for the calorons easily seen to be delta functions [13]. This was already realised long ago by Nahm himself [6], but up to now this has not led to explicit construction of solutions.

While finishing this paper we became aware of Ref. [14] in which some of the same issues are addressed.

2. The solutions

For calorons we compactify the time direction by periodic identification. One requires the gauge fields to be periodic up to a gauge transformation. By a suitable choice of gauge, where $A_0$ tends to zero at infinity, and the topological charge $k$ is realised by the winding number of the gauge transformation that describes $A_i$ at spatial infinity, one has

$$A_\mu(x,x_0 + 1) = \exp(2\pi i \omega \cdot \tau) A_\mu(x,x_0) \times \exp(-2\pi i \omega \cdot \tau), \quad (1)$$

with $\tau_i$ the Pauli matrices. We have chosen units such that the period in the time direction equals one. Its proper value, where relevant, can be reinstated later on dimensional grounds.

The Polyakov loop, $P(x) = \exp(\int_0^1 dt A_0(x))$, is seen to satisfy

$$P_\infty \equiv \lim_{|x| \to \infty} P(x) = \exp(2\pi i \omega \cdot \tau). \quad (2)$$

For the periodic case ($\omega = |\omega| = 0$) the caloron solutions are well known [1], but to this date no solutions
for the general case were known. Although it was argued that for the case of non-trivial values of $P_n$ these solutions are not important in the finite temperature partition function [15], it might be worthwhile to reinvestigate this issue now that the solutions are known explicitly. As the finite temperature partition function requires the physical, i.e. gauge invariant, components of the fields to be periodic, we do in principle have to include also the configurations with $P_n$ non-trivial.

We will first give the explicit solution before discussing its construction. Using a rotation, we can achieve $\hat{\omega} \cdot \tau = \omega \cdot \tau / \omega = \tau_3$, with $\omega \in [0, \frac{1}{2}]$. The solution is written in terms of one real ($\phi(x)$) and one complex ($\chi(x)$) function, and in terms of the (anti-) self dual 't Hooft tensors [16] $(\bar{\eta}_{\mu \nu}^i)$ (with our conventions of $t = x_0$, $e_{0123} = -1$)

$$
\eta_{00}^i = -\eta_{0j}^i = \bar{\eta}_{ij}^j = -\bar{\eta}_{00}^j = \delta_{ij}, \quad \eta_{jk}^i = \bar{\eta}_{jk}^i = e_{ijk}.
$$

We find

$$
A_\mu(x) = i \frac{1}{2} \tau_3 \bar{\eta}_{\mu \nu}^i \partial_{\nu} \log \phi + i \frac{1}{2} \text{Re}\left((\tau_i + i \tau_2)(\bar{\eta}_{\mu \nu}^i - \bar{\eta}_{\nu \mu}^i)\partial_{\nu} \chi\right) \phi,
$$

where

$$
\phi = \psi / \hat{\psi}, \quad \chi = \pi \rho^2 \psi^{-1} e^{i \pi / \omega x_0} \times \left(s^{-1} \sinh(4 \pi s \omega) e^{-2 \pi i x_0} + r^{-1} \sinh(4 \pi r \bar{\omega})\right), \quad \bar{\omega} = \frac{i}{2}(1 - 2 \omega), \quad \hat{\psi} = \cos(4 \pi s \omega) \cos(4 \pi r \bar{\omega}) + \frac{(r^2 + s^2 - \pi \rho^2)}{2rs} \sinh(4 \pi s \omega) \sinh(4 \pi r \bar{\omega}) - \cos(2 \pi x_0), \quad \psi = \hat{\psi} + \pi \rho^2 (s^{-1} \sinh(4 \pi s \omega) \cosh(4 \pi r \bar{\omega})) + r^{-1} \sinh(4 \pi r \bar{\omega}) \cosh(4 \pi s \omega) + \frac{\pi \rho^4}{rs} \sinh(4 \pi s \omega) \sinh(4 \pi r \bar{\omega}).
$$

The two radii, $r$ and $s$, that appear in Eq. (5) are defined by

$$
r^2 = (x + 2 \pi \rho^2 a)^2, \quad s^3 = (x - 2 \pi \bar{\omega} \rho^2 a)^2, \quad a = \hat{\omega}.
$$

and in a sense the solution can be seen as being built from a suitable combination of two dyons (BPS monopoles) of opposite charge, best understood in terms of an old construction by Taubes involving non-contractible loops in Yang-Mills configuration space [17]. The parameter $\rho$ is related to the scale of the instanton solution, and the two constituent BPS monopoles are separated by a distance $\pi \rho^2$. Their mass ratio approaches $\omega / \bar{\omega}$ for large $\rho$, when the solution becomes static (in a suitable gauge). Some of these features are illustrated in Fig. 1. For $\omega = 0 \equiv 0 \mod \frac{1}{2}$, $P_n = \pm 1$, the gauge symmetry is no longer broken to the $U(1)$ subgroup generated by $\hat{\omega} \cdot \tau$. In this case one of the two radii will drop out of the problem and the solution is spherically symmetric. The relation to the BPS monopole for large $\rho$ at $\omega = 0$ can already be found in Ref. [18]. The constituent monopole description is also the basis for the results in Ref. [14], and seems to be the natural framework for discussing the situation for arbitrary gauge groups, going back to the work of Nahm [6].

For the case $P_n = \pm 1$ one finds $\chi = \chi^* = 1 - \phi^{-1}$ and $A_\mu = \frac{i}{2} r \tau_3 \bar{\eta}_{\mu \nu}^i \partial_{\nu} \log \phi$, which is in the form of the celebrated 't Hooft ansatz [19], and for which $\phi^{-1} \bar{\eta}^i \partial^i \phi = 0$. For non-trivial values of $\omega$ such a simple characterisation is not readily available. Nevertheless, the expression of $\text{tr} F_{\mu \nu}^2(x) = -\partial^2 \bar{\eta}_{\mu \nu}^i \partial^i \log \phi$, derived for the 't Hooft ansatz [19], has a remarkable generalisation to the case of non-trivial $\omega$.

$$
\text{tr} F_{\mu \nu}^2(x) = -\partial^2 \bar{\eta}_{\mu \nu}^i \partial^i \log \psi.
$$

This equation was used for constructing Fig. 1. We have also computed numerically the curvature directly from Eq. (4), checking the self-duality and verifying Eq. (7).

Finally we note that $\omega$ should not be considered as part of the moduli. For each $\omega$ one has a different set of solutions. This is particularly clear when we transform to the periodic gauge, for which $A_0 = 2 \pi i \omega \cdot \tau$ at $|x| \to \infty$. For each choice of $\omega$ we have
Fig. 1. Profiles for calorons at $v = 0, 0.125, 0.25$ from top to bottom with $r_s$. The axis connecting the lumps, separated by a distance $p$, corresponds to the direction of $v$. The other direction indicates the distance to this axis, making use of the axial symmetry of the solutions. Vertically is plotted the action density, at the time of its maximal value, on equal logarithmic scales for the three profiles. The profiles were cut off at an action density below $v_{\text{re}}$. The mass ratio of the two lumps is approximately $v_r v$, i.e. zero no second lump, a third and one equal masses, for the respective values of $v$.

an eight dimensional moduli space with as parameters the position of the caloron, which can be obtained by translating our solution in space and time, the scale $\rho$ and a combined rotation and gauge transformation (keeping $P_\infty$ fixed).

3. The construction

Rather than presenting the Nahm transformation [5,6], we make a shortcut by describing the ADHM construction [10], and showing how from an infinite periodic array of instantons (no longer oriented parallel in group space [15]) we can obtain the caloron. Input from the Nahm transformation comes at the point where we interpret the quaternionic valued matrices and vectors that appear in the ADHM construction (collectively known as ADHM data) as the Fourier coefficients with respect to $z$, appearing in the Nahm transformation. The main advantage of this approach is that already much is known about the calculus of multi-instantons in the ADHM formalism [11], in particular also for computing the metric on the moduli space [20]. The ADHM data are obtained after applying the Nahm transformation. Applying this transformation for the second time yields the construction of the self-dual field in terms of the ADHM parameters [13].

Specifically, for charge $k$ SU(2) instantons the ADHM data are given by a quaternionic valued vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ and a symmetric quaternionic valued $k \times k$ matrix $B$. We parametrise the quaternions as linear combinations of the unit quaternions, $\sigma_0 = 1$ and $\sigma_i = i \sigma_i$. The vector $\lambda$ is directly related to the asymptotic behaviour of the zero-modes for the Weyl operator, which gives rise to the boundary terms mentioned in the introduction (see Ref. [6,11,13] for details). This will be seen to be responsible for the announced delta function singularities in the case of calorons. The matrix $B$ is related directly to the connection for the Nahm bundle. In order for the ADHM data to describe a self-dual connection, they have to satisfy a quadratic relation which states that $B^2 B + \lambda^2 \lambda$ is a non-singular symmetric $k \times k$ matrix whose entries are real (i.e. proportional to $\sigma_0$). Alternatively one may state that $B^2 B + \lambda^2 \lambda$ has to commute with the quaternions.

We replace $B$ by $B - x$, with $x = x_\mu \sigma_\mu$ (a $k \times k$ unit matrix is implicit in our notation). The quadratic ADHM relation obviously remains valid. We note that $x$ corresponds precisely to adding the flat $U(1)$ connection to the Nahm connection when applying the Nahm transformation for the second time. The self-dual gauge field is now given by [10]

$$A_\mu(x) = \frac{u'(x)(\partial_\mu u(x)) - (\partial_\mu u'(x))u(x)}{2(1 + u'(x)u(x))},$$

$$u'(x) = \lambda(B - x)^{-1}.$$ (8)
There remains a redundancy which can be related to the gauge invariance for the Nahm bundle,
\[ \lambda \rightarrow q \lambda T, \quad B \rightarrow T^{-1}BT \], \quad \lambda_\mu(x) \rightarrow q A_\mu(x) \tilde{q}, \tag{9} 
where \( q \) is a unit quaternion \( (q \tilde{q} = |q|^2 = 1) \), i.e.

a constant gauge transformation (we denote \( \tilde{x} \) to denote the conjugate quaternion \( x^\dagger \)), note \( \tilde{q} = q^{-1} \), and \( T \) is an orthogonal \( k \times k \) matrix with real entries. This can be used to count the number of moduli of a charge \( k \) instanton, being \( 8k - 3 \). Including the \( q \) as moduli gives \( 8k \) parameters, forming a hyperKähler manifold [12,21].

The boundary condition \( A_\mu(x) = \exp(2\pi i \omega \cdot \tau) A_\mu(x) \exp(-2\pi i \omega \cdot \tau) \) is compatible with the algebraic nature of the ADHM construction, and can be implemented by

\[ \lambda_n = \exp(2\pi i \omega \cdot \tau) \zeta, \quad B_{m,n} = B_{m-1,n-1} + \delta_{m,n}, \tag{10} \]

with \( \zeta \) an arbitrary quaternion, such that

\[ u_{m+1}(x+1) = u_m(x) \exp(-2\pi i \omega \cdot \tau). \tag{11} \]

We note that this means \( k = \infty \). Indeed, \( A_\mu(x) \) viewed as a solution on \( R^4 \) with unit topological charge per period has an infinite total topological charge. For trivial holonomy \( (\omega = |\omega| = 0 \mod \frac{1}{2}) \) it is seen that the quadratic constraint on the ADHM data is solved by choosing \( B_{m,n} = (m + \xi) \delta_{m,n} \), with \( \xi \) an arbitrary quaternion, which describes the position of the caloron. The caloron size is given by \( \rho = |\xi| \) and \( \xi/\rho \) represents a constant gauge transformation.

The major obstacle for non-trivial holonomy was satisfying the non-linear constraint. This can be solved most easily by introducing a Fourier transformation [22]. Let us first give the solution in the matrix representation

\[ B_{m,n} = (m + \xi) \delta_{m,n} + \lambda \delta_{m,n}, \]

\[ \lambda_{m,n} = i \tilde{\omega} \cdot \tau \frac{\sin(2\pi \omega (m-n))}{m-n}(1 - \delta_{m,n}). \tag{12} \]

It has the right number of parameters, \( 8 \) in total, where \( q \equiv \xi/\rho \) is split in a \( U(1) \) part commuting with \( P_n \), describing the residual \( U(1) \) gauge invariance, and a part \( SU(2)/U(1) \) describing a rotation of the vector \( \omega \), compensated by a gauge transformation to ensure that \( P_n \), or equivalently the periodicity condition, is unaltered.

It is advantageous to first give some general results, valid for arbitrary instantons, giving a more efficient way of representing \( A_\mu(x) \). The derivation is straightforward and will be given elsewhere. It is well known that in this problem two Green’s functions appear [6,11]. One is associated to the quadratic ADHM relation

\[ f_s = (\Delta^T(x) \Delta(x))^{-1}, \quad \Delta^T(x) \Delta(x) = (B - x)^{(B - x)^T} + \lambda \Delta. \tag{13} \]

\[ \Delta^T(x) \Delta(x) = (B - x)^{(B - x)^T} + \lambda \Delta. \tag{14} \]

From the definition of \( u(x) \) it follows that

\[ \phi = 1 + \lambda G_s \lambda^T = 1 + u^T(x) u(x). \tag{15} \]

One finds the following compact result

\[ A_\mu(x) = - \frac{1}{2} \phi \delta \left( \phi^{-1} \left( \lambda \bar{\eta}_{\mu s} G_s \lambda^T \right) \right) \], \tag{16} 

where \( \bar{\eta}_{\mu s} \equiv \sigma_3 \eta_{\mu s} \) (the ’t Hooft tensors [16] may be defined through \( \bar{\eta}_{\mu s} = \frac{i}{2}(\sigma_3 \sigma_\mu \sigma_\nu - \sigma_\mu \sigma_3 \sigma_\nu) \) and \( \eta_{\mu s} = \frac{i}{2}(\sigma_\mu \sigma_\nu - \sigma_\nu \sigma_\mu) \)).

By choosing \( B \) diagonal and the entries of \( \lambda \) real, this result immediately leads to the subclass of solutions that are expressed in terms of the ’t Hooft ansatz [19]. The quadratic ADHM condition is obviously satisfied, and \( A_\mu(x) = \frac{1}{2} \bar{\eta}_{\mu s} \delta \log \phi \).

We note that the Green’s functions \( f_s \) and \( G_s \) are intimately related. In particular

\[ G_s \lambda^T = \phi f_s \lambda^T \], \tag{17} 

which implies that (see also Ref. [23]; their conventions relate to ours by \( u \rightarrow \lambda^T, \phi \rightarrow \phi^{-1} \))

\[ \phi = (1 - \lambda f_s \lambda^T)^{-1}, \quad A_\mu(x) = - \frac{1}{2} \phi \delta \left( \lambda \bar{\eta}_{\mu s} f_s \lambda^T \right). \tag{18} \]

As a consequence we will only need to know \( f_s \).

This Green’s function is simpler to determine than
G, since \( f_s \) is proportional to \( \sigma_l \). A standard computation, that lies at the heart of showing that the curvature obtained after applying the Nahm transformation is self-dual \([6,11,8]\) (for which it is crucial that \( f_s \) commutes with the quaternions), yields

\[
F_{\mu\nu} = 2\phi^{-1} u\eta_{\mu\nu} f_s u. \tag{19}
\]

We have now reduced the explicit computation of instanton solutions to the computation of \( f_s \). Incidentally, it can be verified that all infinite sums involved in expressions that appear for the calorons are convergent. Eq. (19) demonstrates that it gives a self-dual solution. We now discuss the Fourier transformation, in terms of which one solves for the quadratic ADHM constraint and for the Green’s function \( f_s \). One defines

\[
\hat{A}(z) = \sum_m \exp(2\pi i m z) A_m = (P_+ \delta(z - \omega) + P_- \delta(z + \omega)) \xi,
\]

\[
\delta(z - \omega) \hat{D}(z) = \sum_m \exp(2\pi i (mz' - nz)) B_{m,n}.
\]

(20)

where \( P_\pm = \frac{1}{2}(1 \pm \hat{\omega} \cdot \tau) \). Parametrising \( B_{m,n} \) as before in terms of \( \xi \) and \( \hat{A}_{m,n} \), with \( \delta(z - \omega) \hat{A}(z) = \sum_m \exp(2\pi i m z) \hat{A}_{m,n} \), we find

\[
\hat{D}(z) = \frac{1}{2\pi i} \frac{d}{dz} + \xi + \hat{A}(z).
\]

Thus, \( B \) has been turned into a differential operator, precisely the Weyl operator appearing in the Nahm transformation, with \( \hat{A}(z) \equiv A_{m,n}(z) \sigma_m \) the connection for the Nahm bundle \([13]\) (up to factors \( 2\pi i \), to match with the conventions of the ADHM construction). The Nahm transformation would require \( \hat{D}^2 \) to commute with the quaternions, which is equivalent to saying that the curvature of the Nahm connection is self-dual. Due to the boundary terms discussed in the introduction, this self-duality is violated at a finite number of points \([13]\) and the presence of \( \hat{A}^2 \) in the quadratic ADHM relation is precisely so as to correct for the violations of self-duality, in accordance with the expectations expressed in the introduction. After Fourier transformation this quadratic relation reads, with a slight abuse of notation,

\[
\left( \Delta'(x) \Delta(x) \right)(z) = \left( \hat{D}(z) - x \right) \left( \hat{D}(z) - x \right) + \hat{A}(z).
\]

(22)

where

\[
\delta'(z - z') \hat{A}(z) = \hat{A}'(z) \delta(z - z'),
\]

\[
\hat{A}(z) = \xi(z)(P_+ \delta(z - \omega) + P_- \delta(z + \omega)) \xi.
\]

(23)

The condition that \( \Delta'(x) \Delta(x) \) has to commute with the quaternions is now seen to lead to the equation

\[
d\hat{A}(z)/dz = \pi \xi \hat{\omega} \cdot \sigma \xi \left( \delta(z + \omega) - \delta(z - \omega) \right),
\]

(24)

which is solved by (eliminating an arbitrary additive constant that can be absorbed in \( \xi \) by imposing \( f_s/ \xi dA(z) = 0 \)),

\[
\hat{A}(z) = \xi \hat{\omega} \cdot \sigma \xi \hat{A}^{(0)}(z),
\]

(25)

\[
\hat{A}^{(0)}(z) = \pi(1 - 2\omega - x_0(z)),
\]

where \( x_0(z) = 1 \) for \( \omega < z < 1 - \omega \) (requiring \( \omega \in [0, \frac{1}{2}] \)) and 0 elsewhere. Fourier transformation of \( \hat{A}(z) \) yields the result in Eq. (12).

As %x and \( \xi \) always occur in the combination \( x - \xi \), we absorb \( \xi \) by a translation in \( x \). The computation of \( f_s \) now reduces to a one-dimensional quantum mechanical problem on the circle,

\[
\left\{ \left( \frac{1}{2\pi i} \frac{d}{dz} - x_0 \right)^2 + r^2 \chi_\omega(z) + s^2 (1 - \chi_\omega(z)) \right\} f_s(z, z')
\]

(26)

where the radii \( r \) and \( s \) were given in Eq. (6) (note that here \( a \cdot \sigma = \xi \hat{\omega} \cdot \sigma \xi / |\xi|^2 \)). We will present the explicit analytic solution for \( f_s(z, z') \) elsewhere, but it should be noted that, due to the particular form of \( \hat{A}(z) \), only \( f_s(\omega, \omega) = f_s(-\omega, -\omega) \) and \( f_s(\omega, -\omega) = f_s(-\omega, \omega) \), respectively real and complex functions of \( x_\mu \), will occur in the evaluation of \( A_\mu(z) \) (see Eq. (18)). To obtain Eq. (4) from Eq. (18), one moves \( \eta_{\mu\nu} \) through \( \hat{A}(z) \), which for non-trivial \( \omega \) do not commute.
We close this section by quoting a useful and remarkable result [11,20],
\[
\text{tr} F_{\mu}^2(x) = - \delta_1 \delta_2 \text{logdet}(f_i),
\]
which leads to the result given in Eq. (7). Although for the caloron \text{logdet}(f_i) is divergent, \( \delta_1 \text{logdet}(f_i) \) is well defined.

4. The geometry of moduli space

The moduli space of the self-dual solutions is given by the ADHM data, or equivalently by the Nahm connections. The latter are self-dual connections on the dual space and the Nahm transformation provides precisely a T-duality [9]. It has been well established that this transformation preserves the metric and hyperKähler structure of the moduli spaces. Establishing that this transformation preserves the periodicity is well defined. We use results due to Osborn [20], which can be readily transposed to the case of the calorons and become particularly elegant after the Fourier transformation.

Since we have a closed expression for \( A_\mu(x) \) in terms of the ADHM parameters, we can compute the variations \( \delta A_\mu(x) \) with respect to the moduli in terms of variations of the ADHM data, summarised in terms of \( \delta \Delta(x) \). The metric is obtained by computing \( ||\delta A||^2 = -\int d^4 x \text{tr}(P \delta A_\mu(x))^2 \), where \( P \) is the projection on the transverse gauge fields, achieved by applying an infinitesimal gauge transformation such that \( \delta A_\mu(x) \) satisfies the background gauge condition \( D_\mu \delta A_\mu(x) = 0 \). It can be shown that [20]
\[
D_\mu \delta A_\mu(x) \big|_{x = 0} = \delta_1 u^i f_i \sigma_\mu(\delta \Delta^i - \Delta^i(\delta \Delta)) \sigma_\mu f_i u,
\]
which vanishes if and only if \( \frac{1}{2} \text{tr}(\delta \Delta^i - \Delta^i(\delta \Delta)) = 0 \). This condition is precisely the background gauge condition for the Nahm connection (for \( T^4 \) this is an exact statement [8], whereas in general \( \lambda \) provides the corrections due to the asymptotic behaviour of the chiral zero-modes of the Weyl operator). In particular this implies that the background gauge condition is preserved under the Nahm, or T-duality, transformation. Projection to a transverse variation of the connection can therefore be achieved by applying an infinitesimal gauge transformation to the ADHM data as given in Eq. (9) \( (q = 1) \). Under such an infinitesimal gauge transformation, \( T = \exp(\delta X) = 1 + \delta X + \cdots \),
\[
\delta X \lambda = \lambda \delta X, \quad \delta X = [B, \delta X], \quad \delta X^t = -\delta X.
\]
Replaced \( \delta \Delta \) by \( C_x = \delta \Delta + \delta \Delta \), the background gauge condition gives an equation for \( \delta \Delta \) in terms of \( \delta \Delta \),
\[
\frac{1}{2} \text{tr}(B'[B, \delta X] - B'[\delta X]B + 2 \delta X A + \Delta^2 \delta \Delta
\]
\[
- \delta \Delta^2 \Delta = 0.
\]
For the caloron, with \( \delta X \) preserving the periodicity (11), the transformation of \( \delta \Delta \) to a transverse variation can now be reformulated after Fourier transformation as
\[
\frac{1}{4 \pi^2} d^2 \delta \vec{X}(z) + \| \vec{X} \|^2 \theta(z - \omega) + \theta(z + \omega) \delta \vec{X}(z)
\]
\[
= \frac{i}{4} \text{tr}((\delta \vec{X} - \vec{X} \delta \vec{X}) \vec{\omega} \cdot \vec{\sigma})
\]
\[
\times (\delta(z - \omega) - \delta(z + \omega)),
\]
where \( \delta(z' - z) \delta \vec{X}(z) = \sum_{m,n} \exp(2 \pi i (m z' - n z)) \delta X_{m,n} \). Solving for \( \delta \vec{X}(z) \) gives
\[
\delta \vec{X}(z) = -\pi i \frac{\text{tr}((\delta \vec{X} - \vec{X} \delta \vec{X}) \vec{\omega} \cdot \vec{\sigma})}{1 + 4 \pi^2 \omega(1 - 2 \omega)|\vec{\zeta}|^2}
\]
\[
\times \int_0^\zeta d\zeta \vec{\Theta}(0)(z'),
\]
which is a zig-zag function (periodic and odd in \( z \)), with discontinuous derivatives at \( z = \pm \omega \).

The following miraculous formula due to Osborn [20] allows us to compute \( ||\delta A||^2 \),
\[
\text{tr}(P \delta A_\mu(x))^2 = \frac{1}{2} \delta_2^2 \text{tr}(C_x(2 - \Delta(x)) f_i \Delta^i(x) C_x f_i).
\]
Since the right-hand side of this equation is smooth and a total derivative, the integration over space and time is completely determined by the behaviour for \( r = |x| \to \infty \). In this limit we may replace \( f_i(z, z') \) by
\[ \pi r^{-1} \exp(-2\pi |z-z'| + 2\pi i x_0 (z-z')) \] (near \( z = z' \), properly extended as a function on \( S^1 \times S^1 \)). From this we find

\[
\|\delta A\|^2 = 2\pi^2 \text{tr} \int_0^1 dz \left\{ d\tilde{B}^*(z) d\tilde{B}(z) + 2d\tilde{\lambda}^*(z) \int_0^1 d\zeta d\tilde{\lambda}(z') \right\},
\]

where

\[
d\tilde{\lambda}(z) = P_+ \delta(z - \omega)(\delta \zeta + \xi \delta \tilde{X}(\omega)),
\]

\[
d\tilde{\lambda}^*(z) = P_- \delta(z + \omega)(\delta \zeta - \xi \delta \tilde{X}(\omega)),
\]

\[
d\tilde{B}(z) = \delta \xi + \xi \tilde{A}(z) + \frac{1}{2\pi i} \frac{d\delta \tilde{X}(z)}{dz}.
\]

For the metric one finds the explicit result

\[
\|\delta A\|^2 = 4\pi^2 |\delta \xi|^2 + 8\pi^2 (1 + R^2) |\delta \zeta|^2 - 2\pi^2 R^2 |\xi|^2 \left(1 + \frac{1}{1 + R^2} \right)(\hat{\omega} \cdot \delta \sigma)^2.
\]

where \( (\xi \equiv \gamma_\mu \partial_\mu) \)

\[
R^2 = \pi^2 |\xi|^2 / M^2, \quad M^2 = 8\omega(1 - 2\omega),
\]

\[
\frac{1}{\pi} \delta \sigma \gamma_\mu |\xi|^2 \gamma_\mu dy_\nu.
\]

One readily recognises, putting \( \delta \xi = 0 \), the Taub-NUT metric [24] with mass \( M \). For \( \hat{\omega} \) in the third direction this metric is given by [25]

\[
ds^2 = \left(1 + \frac{x^2}{16M^2} \right) \left(dx^2 + \frac{1}{4}x^2 (d\sigma_1^2 + d\sigma_2^2) \right) + \frac{1}{4}x^2 d\sigma_1^2 + \frac{1}{16M^2} \left(1 + \frac{x^2}{16M^2} \right),
\]

where we identify \( x^2 = 8\pi^2 p^2 \). We note that the Taub-NUT space is a self-dual Einstein manifold [26] and that it has a hyperKähler structure [27], inherited from the hyperKähler structure of \( R^1 \times S^1 \).

5. Conclusions

We have found the explicit charge one \( SU(2) \) caloron solutions with the Polyakov loop at spatial infinity non-trivial. Previously only solutions for which the latter was trivial were known [1]. Those

\[ \text{were argued to dominate in the instanton contribution to the finite temperature partition function [15], a question that can now more directly be addressed and is perhaps of physical significance.} \]

We have shown that the moduli space of these solutions forms a Taub-NUT space, providing an exact classical T-duality between H-monopoles and Kaluza-Klein monopoles [4]. Indeed it is well-known that the Taub-NUT metric describes the spatial part of the Kaluza-Klein monopole [28] with compactification radius \( 4M \). Most importantly we have related the holonomy to the compactification radii involved in the dual descriptions.

6. Note added in proof

For clarity we emphasise the obvious fact that the Taub-NUT space is a double cover of the moduli space of framed instantons. The relevant identification, \( \zeta \rightarrow -\zeta \), corresponds to the \( Z_2 \) gauge invariance and leaves the gauge field unaltered. It has the origin as a fixed point (resulting in an orbifold singularity for the moduli space). Indirect arguments concerning the nature of the moduli space can be found in Ref. [2, 29]. See also the results in Ref. [30], which appeared after the completion of our paper.

Acknowledgements

This work was started when one of us (P.v.B.) was visiting the Newton Institute during the first half of 1997. He thanks the staff for their hospitality and gratefully acknowledges discussions with Jerome Gauntlett, Nick Manton, Werner Nahm, Hiraku Nakajima, David Olive and Erik Verlinde. T.C.K. was supported by a grant from the FOM/SWON Association for Mathematical Physics.

References


Geometric actions for D-branes and M-branes

M. Abou Zeid, C.M. Hull

Physics Department, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK

Received 27 February 1998
Editor: P.V. Landshoff

Abstract

New forms of Born-Infeld, D-brane and M theory five-brane actions are found which are quadratic in the abelian field strength. The gauge fields couple both to a background or induced metric and a new auxiliary metric, whose elimination reproduces the non-polynomial Born-Infeld action. This is similar to the introduction of an auxiliary metric to simplify the Nambu-Goto string action. This simplifies the quantisation and dualisation of the gauge fields. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

In string theory, it has proved fruitful to replace the Nambu-Goto action which gives the area of the string worldsheet with a classically equivalent action involving a worldsheet metric and a local conformal symmetry \[ w_{x}^{1+3} \]. The Nambu-Goto action is non-polynomial in the string coordinates, whereas the equivalent action is quadratic in the derivatives of the coordinates, greatly simplifying the analysis and allowing a covariant quantisation \[ 4 \]. This has a generalisation for the Nambu-Goto action for \( p \)-branes (proportional to the world-volume), but the resulting theory is only conformally invariant for the string case, \( p = 1 \). The purpose of this paper is to propose and investigate an action that may play a similar role for the Born-Infeld theory of electromagnetism, and its D-brane generalisations. The Born-Infeld action is non-polynomial in the field strength \( F_{\mu\nu} \), but introducing a new intrinsic auxiliary metric gives a classically equivalent action which is quadratic in \( F_{\mu\nu} \), and which has a classical conformal symmetry in four dimensions, instead of the two dimensions for the \( p \)-brane world-volume. There are similar actions for the generalisations of Born-Infeld theory governing the effective worldvolume theories of D-branes \[ 5–15 \] and M-branes \[ 16–19 \]. As the new actions are quadratic in \( F_{\mu\nu} \), integration over the gauge fields is straightforward and, just as in string theory, the focus turns to the integration over metrics. The new action can be used to dualise the Born-Infeld gauge field in all dimensions, circumventing the problems arising in other approaches. In particular, it promises to be more convenient than the action presented in Ref. \[ 20,21 \] which used an auxiliary tensor field consisting of a metric together with an antisymmetric part.
2. Actions

We begin with the Born-Infeld action in $p + 1$ dimensions [22]

$$S = - T_p \int d^{p+1} \sigma \sqrt{- \text{det}(g_{\mu \nu} + F_{\mu \nu})} ,$$  \hspace{1cm} (1)

where

$$F_{\mu \nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$  \hspace{1cm} (2)

is the field strength of a $U(1)$ gauge field $A_\mu$, $\mu, \nu = 0, \ldots, p$ are space-time indices and $g_{\mu \nu}$ is the space-time metric. We now show that the action (1) can be written in a form which is quadratic in the field strength $F$, and is therefore simpler to analyse and quantise. The key is to use the fact that

$$\text{det}(g_{\mu \nu} + F_{\mu \nu}) = \text{det}(g_{\mu \nu} - F_{\mu \nu})$$  \hspace{1cm} (3)

to write the integrand in (1) in the form

$$[\text{det}(g_{\mu \nu} + F_{\mu \nu})]^\frac{1}{2} = [\text{det}(g_{\mu \nu} + F_{\mu \nu})]^\frac{1}{2} \text{det}(g_{\mu \nu} - F_{\mu \nu})]^\frac{1}{2}$$  \hspace{1cm} (4)

where $g \equiv \text{det}(g_{\mu \nu})$. The action (1) can thus be rewritten as

$$S' = - T_p \int d^{p+1} \sigma (- g)^\frac{1}{2} (- \mathcal{B})^\frac{1}{2},$$  \hspace{1cm} (5)

where

$$\mathcal{B}_{\mu \nu} = g_{\mu \nu} - g^{\rho \sigma} F_{\rho \mu} F_{\rho \nu},$$  \hspace{1cm} (6)

and $\mathcal{B} \equiv \text{det}([\mathcal{B}_{\mu \nu}])$. Introducing an intrinsic metric $\gamma_{\mu \nu}$ allows us to rewrite (5) in the following classically equivalent form which is quadratic in the gauge field strength $F_{\mu \nu}$

$$S' = - T_p \int d^{p+1} \sigma (- g)^\frac{1}{2} (- \gamma)^\frac{1}{2} \left[ \gamma^{\mu \nu} \mathcal{B}_{\mu \nu} - (p - 3) \Lambda \right]$$  \hspace{1cm} (7)

where $\gamma \equiv \text{det}(\gamma_{\mu \nu})$ and $\Lambda$ is a constant. For $p \neq 3$, the $\gamma_{\mu \nu}$ field equation implies

$$\gamma_{\mu \nu} = \frac{1}{\Lambda} \left( g_{\mu \nu} - g^{\rho \sigma} F_{\rho \mu} F_{\rho \nu} \right),$$  \hspace{1cm} (8)

and substituting back into (7) yields the action (5), which is identical to the Born-Infeld action (1). The constants $T_p, T'_p$ are related by

$$T'_p = \frac{4}{\Lambda} A \frac{p-3}{p} T_p,$$  \hspace{1cm} (9)

For $p = 3$, the four-dimensional action (7) is invariant under the Weyl transformation

$$\gamma_{\mu \nu} \rightarrow \omega(\sigma) \gamma_{\mu \nu},$$  \hspace{1cm} (10)
and the $\gamma_{\mu\nu}$ field equation implies
\[ \gamma_{\mu\nu} = \Omega \gamma_{\mu\nu}, \tag{11} \]
for some $\Omega$, which is found by taking traces of both sides; this gives
\[ \gamma^{\rho\sigma} \left( g_{\rho\sigma} - g^{\kappa\delta} F_{\kappa\rho} F_{\delta\sigma} \right) \gamma_{\mu\nu} = 4 \left( g_{\mu\nu} - g^{\rho\sigma} F_{\mu\rho} F_{\nu\sigma} \right). \tag{12} \]
Substituting this back into (7) gives (5), which is identical to the Born-Infeld action (1), so that (1) and (7) are classically equivalent.

This can be generalised to the D-brane kinetic term
\[ S = -T_p \int d^{p+1}x e^{-\phi} \sqrt{-\det(g_{\mu\nu} + \mathcal{F}_{\mu\nu})}, \tag{13} \]
where
\[ \mathcal{F}_{\mu\nu} = F_{\mu\nu} - B_{\mu\nu}, \tag{14} \]
$\phi$, $g_{\mu\nu}$ and $B_{\mu\nu}$ are the pullbacks to the worldvolume of the background dilaton, metric and NS antisymmetric two-form fields and $F = dA$, with $A$ the $U(1)$ world-volume gauge field. This action gives the effective dynamics of the zero-modes of the open strings with ends tethered on a D-brane when $F$ is slowly varying, so that corrections involving $\nabla F$ can be ignored, and has therefore played a central role in recent studies of D-brane dynamics and string theory duality [24]. However, the non-linearity of (13) makes it rather difficult to study. In particular, the action (13) is inconvenient for the purpose of quantisation, and its dualisation has proved rather difficult [8,10,25–28]. It is therefore useful to know classically equivalent, alternative forms of this action which have a more tractable dependence on the spacetime coordinates $X$ or on the field strength $F$.

In Ref. [21], we obtained an alternative form of (13) which is linear in $F$ and quadratic in derivatives of $X$ by introducing an auxiliary worldvolume tensor with both symmetric and antisymmetric parts, and discussed the dualisation of the worldvolume gauge field in this approach [21,23]. Here, we give an alternative form of (13) that is quadratic in $F$.

As before, introducing an intrinsic metric $\gamma_{\mu\nu}$ allows us to rewrite (13) in the classically equivalent form
\[ S' = -T_p \int d^{p+1}x e^{-\phi} \sqrt{-\gamma} \left[ \gamma^{\mu\nu} \gamma_{\mu\nu} - (p - 3) A \right] \]
\[ = -T_p \int d^{p+1}x e^{-\phi} \sqrt{-\gamma} \left[ \gamma^{\mu\nu} \left( g_{\mu\nu} - g^{\sigma\sigma} \mathcal{F}_{\mu\sigma} \mathcal{F}_{\nu\sigma} \right) - (p - 3) A \right], \tag{15} \]
where the tensions $T_p$, $T_p'$ are related as in Eq. (9).

The energy-momentum tensor $T_{\mu\nu}$ can be defined from the form (15) of the D-brane kinetic term by
\[ T_{\mu\nu} = -\frac{1}{T_p} \frac{1}{\sqrt{-\gamma}} \frac{\delta S}{\delta \gamma^{\mu\nu}}, \tag{16} \]
and we find
\[ T_{\mu\nu} = \left( -g \right)^{1/2} \left[ -\frac{1}{2} \gamma_{\mu\nu} \left( g^{\rho\sigma} \left( g_{\rho\sigma} - g^{\kappa\delta} F_{\kappa\rho} F_{\delta\sigma} \right) - (p - 3) A \right) + g_{\mu\nu} - g^{\rho\sigma} F_{\mu\rho} F_{\nu\sigma} \right]. \tag{17} \]
This is traceless (i.e. $\gamma^{\mu\nu} T_{\mu\nu} = 0$) if $p = 3$ as a result of the Weyl invariance (10), and the equation $T_{\mu\nu} = 0$ implies the field equation of the metric (8) or (11).

The low-energy effective action for an open type I string includes the terms given by (13) with $p = 9$, but with $g_{\mu\nu}, B_{\mu\nu}$, the space-time metric and anti-symmetric tensor gauge field (rather than their pull-backs) [26].
and can be rewritten in the equivalent form (15) with \( p = 9 \). The dimensional reduction of the type I string action (13) to \( p + 1 \) dimensions gives the action for a D-p-brane in static gauge (and with vanishing RR gauge fields), with the 9 + 1 vector field \( A \) giving rise to a vector and \( 9 - p \) scalars \( \chi_j \) on reduction. The reduction of the form (15) of the action then gives a useful form of the static-gauge D-p-brane action which is quadratic in \( A, X \).

We now turn to the reduction of (15) from \( 9 + 1 \) to \( p + 1 \) dimensions. We use the notation that hatted quantities are ten-dimensional, so \( \mu = 0, \ldots, 9 \), while \( \mu = 0, \ldots, p \) and \( i = p + 1, \ldots, 9 \). Then the vector field \( A_\mu = (A_\mu, X_j) \) gives a vector and \( 9 - p \) scalars \( \chi_j \). We choose (for simplicity) a flat space-time metric \( \tilde{g}_{\mu \nu} = \tilde{h}_{\mu \nu} \) and vanishing 2-form \( \tilde{B}_{\mu \nu} \) and make the following Ansatz for the metric \( \hat{g}_{\mu \nu} \):

\[
\hat{g}_{\mu \nu} = \begin{pmatrix}
\gamma_{\mu \nu} + C^i_{\mu \nu} \chi_j & C^i_{\mu \nu} \chi_j \\
C^i_{\nu \chi_j} & \gamma_{ij}
\end{pmatrix}.
\]

Then the metric \( \hat{g}_{\mu \nu} \) gives, as usual, a \( p + 1 \)-dimensional metric \( \gamma_{\mu \nu}, \) \( 9 - p \) vector fields \( C^i_{\mu \nu} \) and \( (9 - p)(10 - p)/2 \) scalar fields taking values in the coset \( GL(9 - p, R)/SO(9 - p) \). The inverse of (18) is

\[
\hat{g}^{\mu \nu} = \begin{pmatrix}
\gamma^{\mu \nu} & -C^{i \mu} \\
-C^{i \nu} & \gamma^{ij} + C^i_{\mu \nu} \gamma^{\rho \sigma} C_j^{\rho \sigma}
\end{pmatrix},
\]

and its determinant is

\[
\det \hat{g}_{\mu \nu} = \gamma_{\mu \nu} \det \gamma_{ij}.
\]

Setting \( F_{ij} \equiv 0 \) and \( F_{i \mu} \equiv \partial_\mu X_i \), this gives the following static gauge D-p-brane action which is quadratic in both \( F \) and \( \partial X^i \):

\[
S' = -T_p \int d^{p+1} \sigma e^{-\phi} \left[ -\gamma_{\mu \nu} \partial_\mu \chi_j \partial_\nu \chi_j + \gamma^{\rho \sigma} \left( \partial_\rho \chi_j \partial_\sigma \chi_j + \gamma^{\mu \nu} \partial_\mu F_{\rho \nu} F_{\rho \nu} \right) + 2 \gamma^{\mu \nu} \partial_\mu \partial_\nu \chi_j \chi_j \right].
\]

This quadratic action should be a convenient starting point for the study of D-p-brane dynamics, taking into account the Born-Infeld corrections.

The methods above can also be applied to the M-theory five-brane action [16,17]. In the PST formulation, the kinetic part of the action is

\[
S = -T_5 \int d^6 \sigma \sqrt{-g} \frac{\epsilon_{\mu \nu \rho \sigma \tau} H_{\mu \nu \rho} \partial_\tau a}{g(\partial a)^2} - \frac{T_5}{4} \int d^6 \sigma \frac{1}{(\partial a)^2} \tilde{K} \mu \nu H_{\mu \nu \rho} g^{\rho \sigma} \frac{\partial a}{g(\partial a)^2},
\]

where

\[
\tilde{K} \mu \nu \equiv \frac{1}{2} \epsilon_{\mu \nu \rho \sigma \tau} H_{\rho \sigma} \partial_\tau a,
\]

with \( H_{\mu \nu \rho} = 3 a_{\mu \nu} h_{\rho \mu \nu} \) the field strength of the self-dual two-form tensor gauge field \( h_{\mu \nu} \) propagating on the world-volume, \( a \) denotes the PST scalar [16] and

\[
(\partial a)^2 \equiv g^{\mu \nu} \partial_\mu a \partial_\nu a.
\]

Introducing an intrinsic metric \( \gamma_{\mu \nu} \) as before, the action (22) can be rewritten in the classically equivalent form

\[
S = -T_5 \int d^6 \sigma \left[ -g^{\mu \nu} \partial_\mu \chi_j \partial_\nu \chi_j \gamma_{\mu \nu} + \gamma^{\rho \sigma} \partial_\rho \chi_j \partial_\sigma \chi_j + \gamma^{\mu \nu} \partial_\mu F_{\rho \nu} F_{\rho \nu} \right] - 2 \Lambda
\]

\[
- \frac{T_5}{4} \int d^6 \sigma \frac{1}{(\partial a)^2} \tilde{K} \mu \nu H_{\mu \nu \rho} g^{\rho \sigma} \frac{\partial a}{g(\partial a)^2}.
\]
This is quadratic in the field strength $H$ and thus is more convenient for gauge field quantisation in the background $g_{\mu\nu}$ than (22).

3. Dual actions

The dualisation of the form (7) of the Born-Infeld action can be achieved via the addition of a Lagrange multiplier term imposing Eq. (2). Consider the action

$$S = -T^0_0 \int d^{p+1} \sigma \left\{ (\gamma - 2)^{(p - 3)} \left[ \gamma^{\mu\nu} - g^{\mu\nu} F_{\mu\nu} - 2(p - 3) A \right] + 2(\text{H}^{\mu\nu} - \partial_\mu A_\nu) \right\},$$

(26)

where $\text{H}^{\mu\nu}$ is a tensor density and $F$ is regarded as an independent field. Integrating out $\text{H}^{\mu\nu}$ sets $\partial_\mu A_\nu$ and yields the original action (7). Alternatively, integrating out $A_\mu$ imposes the constraint

$$\partial_\mu \text{H}^{\mu\nu} = 0,$$

(27)

which can be solved in terms of a $(p - 2)$-form $A$,

$$\text{H}^{\mu\nu} = \frac{1}{(p - 1)!} \epsilon^{\mu\nu\rho \gamma_1 \gamma_2 \ldots \gamma_{p-2}} \partial_\rho A_{\gamma_1 \gamma_2 \ldots \gamma_{p-2}},$$

(28)

where $\epsilon^{\mu\nu\rho \gamma_1 \gamma_2 \ldots \gamma_{p-2}}$ is the alternating tensor density. Now $F$ is an auxiliary two-form occuring quadratically in the action and can be integrated out. The field equation for $F_{\mu\nu}$ is

$$(-g)^{(p - 3)}(\gamma - 2)^{\gamma_{\mu\nu}} g^{\mu\nu} + \gamma^{\nu\rho} g^{\mu\nu} F_{\mu\nu} = 2 \text{H}^{\mu\nu},$$

(29)

where $\text{H}^{\mu\nu}$ is given by the solution (28), and the Gaussian integration amounts to solving this for $F_{\mu\nu}$ and substituting the solution $F[g_{\mu\nu}, \gamma_{\mu\nu}, \text{H}^{\mu\nu}]$ in the action (15). This gives the dual action $S[g_{\mu\nu}, \gamma_{\mu\nu}, \text{H}^{\mu\nu}]$. In principle, an equivalent dual action $S[g_{\mu\nu}, \gamma_{\mu\nu}, \text{H}^{\mu\nu}]$ can then be obtained by integrating out the auxiliary metric $\gamma_{\mu\nu}$, but in practice this procedure is difficult to carry out explicitly because of the non-linearity in the worldvolume metric of Eq. (29) and of the action $S[g_{\mu\nu}, \gamma_{\mu\nu}, \text{H}^{\mu\nu}]$. Defining the matrices

$$f_\mu = F_{\mu\rho} g^{\rho\nu}, \quad h_\mu = (\gamma - 2)^{\gamma} g^{\mu\nu} \text{H}^{\mu\nu}, \quad \beta_\mu = 2(g_{\mu\rho} \gamma^{\rho\nu} - \delta_\mu^{\nu}),$$

(30)

the Eq. (29) can be written as

$$f = f + \{ \beta, f \} = (1 + L_\mu) f,$$

(31)

where for any matrices $X, Y$, the operator $L_X$ is defined by

$$L_X Y = \{ X, Y \}.$$

(32)

Then (31) can be inverted to give

$$f = (1 + L_\mu)^{-1} h = (1 - L_\mu + L_\mu^2 - L_\mu^3 + \ldots) h = h - \{ \beta, h \} + \{ \beta, \{ \beta, h \} \} - \{ \beta, \{ \beta, \{ \beta, h \} \} \} + \ldots$$

(33)

Substituting this solution for $F$ back in (26) gives

$$S = -T^0_0 \int d^{p+1} \sigma \left\{ (\gamma - 2)^{(p - 3)} \left[ \gamma^{\mu\nu} g_{\mu\nu} - (p - 3) A \right] + 2(\gamma - 2)^{(p - 3)} \left[ h(1 + L_\mu)^{-1} h \right] \right\},$$

$$= -T^0_0 \int d^{p+1} \sigma \left\{ (\gamma - 2)^{(p - 3)} \left[ \gamma^{\mu\nu} g_{\mu\nu} - (p - 3) A \right] + 2(\gamma - 2)^{(p - 3)} \left[ (\gamma - 2)^{\gamma} \text{H}^{\mu\nu} M_{\mu\rho\nu} \text{H}^{\rho\nu} \right] \right\},$$

(34)
where the tensor $M_{\mu \nu \rho \sigma}$ is defined by
\[
\text{tr} \left[ h (1 + L_\theta)^{-1} h \right] = h^\mu \rho M_{\mu \nu \rho \sigma} H^{\nu \sigma}
\]
where $h^\mu \rho = g^\mu \rho h_\rho$ and is given to lowest orders by
\[
M_{\mu \nu \rho \sigma} = \gamma_{\mu \nu} B_{\rho \sigma} \left[ \delta_{\rho \sigma}^\rho - \Sigma^\rho \rho \Sigma_\sigma^\sigma + \Sigma\mu^\mu \Sigma^\rho_\rho \Sigma^\sigma_\sigma - \Sigma^\rho_\rho \Sigma^\sigma_\sigma \right] + \Sigma^\rho_\rho \Sigma^\sigma_\sigma \Sigma^\delta_\delta + \ldots
\]
where
\[
\Sigma^\mu_\rho = g^\mu_\rho \gamma_\rho,
\]
and $\Sigma^\mu_\rho$ denotes the inverse of the matrix $\Sigma^\mu_\nu$. The auxiliary metric $\gamma_{\mu \nu}$ occurs algebraically and can in principle be eliminated using its equation of motion, giving $\gamma_{\mu \nu}$ as a function of $g_{\mu \nu}$ and $H^{\mu \nu}$. Although this is hard to do explicitly, it can be done perturbatively, giving $\gamma_{\mu \nu}$ to any desired order in $H^{\mu \nu}$.

The dualisation of the action (15), which is classically equivalent to the D-brane kinetic term (13), proceeds in a similar way. Consider the action
\[
S = -T_\phi d^p+ \sqrt{e^{-\phi} (-g)^2 (-\gamma)^2 \left[ g_{\mu \nu} - g^{\rho \sigma} \mathcal{F}_{\mu \nu} \mathcal{F}_{\rho \sigma} \right] + \left( p - 3 \right) A}
\]
\[
+ 2 \tilde{H}^{\mu \nu} \left( F_{\mu \nu} - 2 \tilde{\eta}_{\mu} A_{\nu} \right) \right). \quad (38)
\]
Integrating out $\tilde{H}^{\mu \nu}$ yields the original action (15). Alternatively, integrating over $A_\mu$ imposes the constraint (27), which is solved in terms of a $(p - 3)$ form $A$ as in (28). Now $F$ is an auxiliary two-form occurring algebraically. The field equation for $F_{\mu \nu}$ is
\[
(-g)^2 (-\gamma)^2 \left[ g_{\mu \nu} \gamma^{\rho \sigma} + \gamma^{\rho \sigma} g_{\mu \nu} \right] \left( F_{\rho \sigma} - B_{\rho \sigma} \right) = 2 \tilde{H}^{\mu \nu}, \quad (39)
\]
where $\tilde{H}^{\mu \nu}$ is given by the solution (28).

Defining the matrix
\[
\hat{f}_\mu = \left( F_{\mu \nu} - B_{\mu \nu} \right) g^{\rho \sigma}, \quad (40)
\]
the Eq. (39) can be written as
\[
h = \tilde{f} + \left\{ \beta, \tilde{f} \right\} = (1 + L_\theta) \tilde{f}, \quad (41)
\]
where the matrices $h$, $\beta$ and the operator $L_\theta$ are defined as in (30) and (32). This can be inverted to give
\[
\tilde{f} = (1 + L_\theta)^{-1} h - \left\{ \beta, h \right\} + \left\{ \beta, \left\{ \beta, h \right\} \right\} + \ldots. \quad (42)
\]
Substituting this solution for $\mathcal{F}$ back in (26) gives
\[
S = -T_\phi d^p+ \sqrt{e^{-\phi} (-g)^2 (-\gamma)^2 \left[ g_{\mu \nu} - \left( p - 3 \right) A \right] + 2 \tilde{H}^{\mu \nu} B_{\mu \nu}}
\]
\[
+ 2 (g)^{-2} (-\gamma)^{-2} \tilde{H}^{\mu \nu} M_{\mu \nu \rho \sigma} \tilde{H}^{\rho \sigma}, \quad (43)
\]
with the tensor $M_{\mu \nu \rho \sigma}$ defined as in (36).

4. Conclusion

In this paper, we have presented new forms of Born-Infeld as well as D-brane and M theory five-brane kinetic terms which are quadratic in the abelian gauge field strength. The gauge fields couple both to a
background or induced metric $g_{\mu\nu}$ and to a new intrinsic metric $g_{\mu\nu}$, and both of these world-volume metrics appear in the action in a remarkably symmetric way. These actions could play an important role in the quantisation of Born-Infeld theory and of the static gauge effective world-volume theories of D-Branes and M-Branes, similar to the role played in string theory by the actions of Ref. [1–4].

The dualisation of the $U(1)$ gauge fields is achieved by adding a Lagrange term imposing the constraint (2), and the dual action is quadratic in the field strength of the appropriate dual potential. The dual action involves an infinite power series in the auxiliary intrinsic metric, which can be eliminated perturbatively.

The four dimensional action (7) with $p = 3$ has a classical Weyl invariance (10), which is closely related to that of the string. Quantum mechanically, this will be anomalous [29,30]. We hope to return to a discussion of this anomaly elsewhere, but we note here that it is trivial to generalise our action to that for a theory of $N$ abelian vector fields and $Nd$ scalars $X_i$, and it is intriguing that it may be possible to choose the numbers $N,d$ to take critical values that give a cancellation of the conformal anomaly, generalising the critical dimension of string theory.

References


\documentclass{article}

\usepackage{amsmath,amssymb}

\begin{document}

\title{p-brane dyons, $\theta$-terms and dimensional reduction}

\author{S. Deser \textsuperscript{a,1}, M. Henneaux \textsuperscript{b,c,2}, A. Schwimmer \textsuperscript{d,3}}

\address{	extsuperscript{a} Department of Physics, Brandeis University, Waltham, MA 02254, USA}
\address{	extsuperscript{b} Faculté des Sciences, Université Libre de Bruxelles, Campus Plaine C.P. 231, B–1050 Bruxelles, Belgium}
\address{	extsuperscript{c} Centro de Estudios Científicos de Santiago, Casilla 16443, Santiago 9, Chile}
\address{	extsuperscript{d} Department of the Physics of Complex Systems, Weizmann Institute, Rehovot 76100, Israel}

\received{11 March 1998}

\editor{H. Georgi}

\begin{abstract}

We present two novel derivations of the recently established $(-)^p$ factor in the charge quantization condition for $p$-brane dyon sources in spacetime dimension $D = 2p + 2$. The first requires consistency of the condition under the charge shifts produced by generalized $u$-terms. The second traces the sign difference between adjoining dimensions to compactification effects.

\end{abstract}

\section{Introduction}

It was recently established\cite{1} that the generalized dyon quantization condition, for $(p-1)$-brane dyons coupled to Abelian $p$-forms in spacetime dimension $D = 2p + 2$, involves a $p$-dependent sign:

$$e\gamma + \left(-\right)^p e_\gamma = 2\pi n\hbar, \quad n \in \mathbb{Z}.\tag{1}$$

The $-$ sign is of course that familiar in $D = 4$ electrodynamics\cite{2–4}. This sign dependence was actually anticipated\cite{5} through analysis of chiral sources coupled to chiral $2p$-forms. It was particularly stressed in\cite{6}, where it was related to supergravity duality groups in higher dimensions\cite{7}. Another approach is based on dyon-dyon scattering; the relation between $D = 10$ and $D = 4$ is also discussed there\cite{8}.

\section{\theta-terms}

a) In $D = 4$ it is well known that adding a $\theta$-term, $(\theta/2)F_{\mu\nu} \ast F^{\mu\nu}$ (always represents dualization) to the Lagrangian has the effect of shifting the electric charge of an $(e, g)$ dyon according to\cite{9}

$$e' = e - 2g\theta.\tag{2}$$

A remarkable feature of this shift is its compatibility with the usual Dirac quantization condition for electric and magnetic charges. Namely, if one simultane-

\end{document}
ously shifts all dyon electric charges according to (2) starting from values \((e_a g_a)\) that obey (1), then the charges \(e_a g_a\) also do, because

\[
e_a g_b - e_b g_a = (e_a - g_a \theta) g_b - (e_b - g_b \theta) g_a = e_a g_b - e_b g_a.
\]

(3)

The \(-\) sign is crucial in this result. Indeed, it is the answer to the converse question: what sign in the quantization condition (1) leaves it invariant under the shift (2)?

b) In \(D = 6\), there is no \(\theta\)-term for a single 2-form since \(F_{ABC} * F^{ABC} \) vanishes identically. However, a \(\theta\)-term is possible with two 2-forms \(A(i), i = 1,2\). The sources here are strings characterized by four strengths ("charges") \((e_a^{(i)}, g_a^{(i)})\), the respective electric (magnetic) charges of string \(a\) coupled to \(A(i)\). We use a uniform convention for the signs of the couplings \((e_a^{(i)}, g_a^{(i)})\) to the 2-forms: the electric couplings enter with the same sign in the minimal coupling term \(\sum_{i=1}^2 e_a^{(i)} A(i)\), for example. Single-valuedness of the wave function leads to the quantization condition

\[
(e_a^{(1)} g_b^{(1)} \pm g_a^{(1)} e_b^{(1)}) + (e_a^{(2)} g_b^{(2)} \pm g_a^{(2)} e_b^{(2)}) = 2 \pi n h, \ n \in \mathbb{Z}
\]

(4)

with a relative \(\pm\) sign between the contributions associated with the two 2-forms because of our identical coupling conventions. We have left the \(\pm\) sign open in (4) to show next how the \(\theta\)-angle argument selects the \(\pm\) sign. [Of course, for sources that couple to only one of the fields -- say \(A^{(1)}\) --, the second term on the left is absent in (4).] Now add the extended \(\theta\)-term

\[
\frac{1}{2} \theta e_{ij} F^{(i)}_{ABC} * F^{(j)ABC}
\]

(5)

to the free Lagrangian \(F_{ABC}^2\). As in \(D = 4\), the effect of this term is to shift the electric charges, but this time by

\[
e_a^{(i)} = e_a^{(i)} - (31) \theta e^{(i)j} g_a^{(j)}.
\]

(6)

The antisymmetry of this shift is traceable to that of the \(\theta\)-term; more explicitly, the \(\theta\)-term involves only mixed couplings, with opposite signs: \(\theta A_{[ij]}^0 g_a B^{(i)m}_{(a)} - \theta A_{[ij]}^0 g_a B^{(j)m}_{(a)}\). [Had we taken opposite conventions for the couplings, there would be a relative minus sign in (4) between the contributions of the two fields, and the same sign in (6).] For the quantization condition (4) to be invariant under the shift (6) then requires the \(\pm\) sign there. Thus, also in \(D = 6\) a (generalized) \(\theta\)-angle argument determines the sign in the quantization condition.

From these two examples, it is clear that the \((-)^p\) factor is a reflection of the opposite symmetries of the \(\theta\)-terms \(F^* F\) and \(e_{ij} F^{(i)} * F^{(j)}\) in alternating dimensions.

3. Adjoining dimensions

We now turn to the argument from dimensional reduction (actually, "enhancement"). Since higher dimension is clearly more restrictive, our logic will be to show that the quantization rule in \(D = 2p - 2\) for those specific configurations obtained by reduction from \(2p\) imposes the form of the \(D = 2p\) rule as well. Specifically we shall show that if the quantization condition holds with one sign in \(D = 2p - 2\), then it must hold with the opposite one in \(D = 2p\); in particular, the \(\pm\) sign in \(D = 6\) follows from the \(-\) sign in \(D = 4\).

We relate \(D = 6\) to \(D = 4\) by toroidal compactification, \(M^6 = R^4 \times T^2\). The spacetime coordinates \(x^A (A = 0,1,\ldots,5)\) split into \(x^4 (x^4, x^5)\) where \((x^4, x^5)\) parametrize the torus which, for our purposes, may be assumed to be the standard \((dx^4)^2 + (dx^5)^2\), with \((x^4, x^5)\) having respective ranges \([0, L_4]\) and \([0, L_5]\). [For fields independent of \((x^4, x^5)\) as considered here, one may always diagonalize the internal metric, but we chose \(not\) to also rescale the ranges to unity.] The full spatial 5\(D\) rotational symmetry is broken by the compactification of course. However, there is a useful residual "\(Y\)-symmetry," under simultaneous interchange of \(x_4/L_4\) with \(x_5/L_5\) together with a 4\(D\) parity (P) transformation. Indeed, we will conclude generally that the quantization
condition in $D = 2p - 2$, together with $Y$-symmetry, implies the corresponding one at $D = 2 p$.

A non-chiral 2-form $A_{\mu \rho}$ in $D = 6$ induces two $D = 4$ U(1) gauge fields $A_\mu^{(i)}$. The reduction proceeds by assuming $A_{\mu \rho}$ to be constant along the internal torus and to have only $A_{4 \mu}$ and $A_{5 \mu}$ as non-zero components. [The other, $A_{\mu \rho}$ components and the higher modes induce further four-dimensional fields which are not relevant to our discussion.]

The correspondence is

$$A_\mu^{(1)} = \frac{1}{L_4 L_5} A_{4 \mu}, \quad A_\mu^{(2)} = \frac{1}{L_4 L_5} A_{5 \mu}$$

as follows from reduction of the 2-form action $\int d^6 x F_{\alpha \beta \rho \sigma \kappa \lambda}$. In $D = 4$ terms, (besides the P) Y-transformations interchange the two $A^{(i)}$.

The $D = 6$ sources that correspond to point particles in $D = 4$ are strings winding around the internal torus directions. For a single string along $x^4$ at $x = 0$, $x^5 = a$, the current has as its only non-vanishing components

$$J^{(4)}_e = e \delta^{(3)}(x) \delta(x^5 - a), \quad J^{(4)}_m = g \delta^{(3)}(x) \delta(x^5 - a)$$

where ($e, g$) are the respective electric and magnetic strengths of the string. The zero modes of the 2-form field couple only to the zero modes of the source. Thus, from the point of view of the zero modes, one can replace the source by a continuous distribution of parallel strings aligned along $x^4$, with constant electric and magnetic strengths per unit length, $(\rho_s, \sigma_s)$, along the transverse $(x^5)$ direction. Such a distribution yields a membrane wrapping around the torus and does not excite the higher modes ("vertical reduction" of [12]). This alternative description preserves translation invariance along $x^5$. Replacing the above source by a stack of strings at $x = 0$ aligned along $x^4$ amounts to replacing the currents of (8) by

$$J^{(4)}_e = \rho_s \delta^{(3)}(x), \quad J^{(4)}_m = \sigma_s \delta^{(3)}(x).$$

These currents are obtained by summing the currents of the individual strings, e.g., $J^{(4)}_e(x, x^5) = \rho_s da \delta^{(3)}(x) \delta(x^5 - a)$ for the string located at $x^5 = a$. The corresponding $D = 6$ charges are

$$e = \rho_s L_5, \quad g = \sigma_s L_5.$$  

From the $4D$ point of view, the stack appears to have the $U(1)$ charges

$$\left( e^{(1)}, g^{(1)}, e^{(2)}, g^{(2)} \right) = \left( \frac{e}{\sqrt{L_4 L_5}}, \frac{g}{\sqrt{L_4 L_5}}, 0, 0 \right)$$

as shown by the analysis of the equations of motion given below. Again, we adopt the same sign conventions for the two $U(1)$'s and define electric and magnetic charges in $4D$ in such a way that $\nabla \cdot E^{(i)} \sim + e^{(i)}$, $\nabla \cdot B^{(i)} \sim + g^{(i)}$ (with same + sign for both $i$). Similarly, the current of a single dyonic string $(e', g')$ lying on the $x^5$ axis at $(x = b, x^5 = c)$ is given by

$$J^{(5)}_e = e' \delta^{(3)}(x - b) \delta(x^5 - c), \quad J^{(5)}_m = g' \delta^{(3)}(x - b) \delta(x^5 - c).$$

Again, from the zero mode point of view, this can be replaced by a suitable stack of strings whose $6D$ currents are

$$J^{(5)}_e = \rho_s \delta^{(3)}(x - b), \quad J^{(5)}_m = \sigma_s \delta^{(3)}(x - b).$$

Here the $D = 6$ charges are

$$e' = \rho_s L_4, \quad g' = \sigma_s L_4,$$

while the $4D$ charges are

$$\left( e^{(1)}, g^{(1)}, e^{(2)}, g^{(2)} \right) = \left( 0, -g' \sqrt{\frac{L_5}{L_4}}, e' \sqrt{\frac{L_5}{L_4}}, 0 \right).$$

These charges have two properties: First, a dyonic string in $D = 6$ along $x^4$ or $x^5$ does not appear as a dyon in $D = 4$. Rather, it is electrically charged for one $U(1)$ and magnetically charged for the other $U(1)$. To get dyons for the same $U(1)$ in $D = 4$, one needs to superpose strings along both $x^4$ and $x^5$.

The same remark applies to the chiral case (in $D = 6$), for which the two $D = 4 U(1)$'s are related by the duality rotation $B^{(2)} = E^{(1)}$, $E^{(2)} = -B^{(1)}$. The above strings would appear as either purely electric (first string) or purely magnetic (second string) but do not carry both types of charges. Second, there is a crucial flip of sign in the magnetic charges for the two $U(1)$'s. The simultaneous existence of the "dual" configurations (15) and (21) reflects the $Y$ symmetry. Indeed, a $Y$-transformation,
to a $D = 4$ observer, just induces this dual exchange. To understand how the $D = 4$ assignments arise, consider the field equations, 
$$\partial_{i}F^{ABC} = J^{BC}_{m},$$ 
$$\partial_{i}F^{ABC} = J^{BC}_{m},$$ for the given sources in terms of the $D = 4$ fields. For the source (10) along $x^4$, the equations reduce to
$$\partial_{i}F^{(1)} = +\sqrt{L_{1}/L_{5}}\partial_{i}F^{04} = +\epsilon^{\sqrt{L_{1}/L_{5}}\delta^{(3)}(x)} \quad (16)$$
and
$$\partial_{i}B^{(2)} = +\sqrt{L_{4}/L_{5}}\partial_{i}(1/2)!\epsilon^{\sqrt{054}F_{mn}}$$
$$= +\sqrt{L_{4}/L_{5}}\partial_{i}F^{05} = +g\sqrt{L_{4}/L_{5}}\delta^{(3)}(x), \quad (17)$$
where $i,m,n = 1,2,3$. For the source (14), one finds
$$\partial_{i}E^{(2)} = \sqrt{L_{4}/L_{5}}\partial_{i}F^{05} = +\epsilon^{\sqrt{L_{4}/L_{5}}\delta^{(3)}(x-b)} \quad (18)$$
and
$$\partial_{i}B^{(1)} = +\sqrt{L_{4}/L_{5}}\partial_{i}(1/2)!\epsilon^{\sqrt{045}F_{mn}}$$
$$= -g\sqrt{L_{4}/L_{5}}\delta^{(3)}(x-b), \quad (19)$$
with a minus sign because $\epsilon^{\sqrt{045}} = -\epsilon^{\sqrt{054}}$. This leads to the assignments (11) and (15).

We now deduce the quantization condition in $D = 6$ from that in $D = 4$. For the strings (11) and (15), the $D = 6$ quantization condition is
$$eg' \pm e'g = 2\pi hn, \ n \in Z. \quad (20)$$
where we have again left the relative sign open. The quantization condition in $D = 4$, on the other hand, is, in terms of $D = 4$ charges,
$$\left(\epsilon_{a}^{(1)}8_{b}^{(1)} - \epsilon_{b}^{(1)}8_{a}^{(1)}\right) + \left(\epsilon_{a}^{(2)}8_{b}^{(2)} - \epsilon_{b}^{(2)}8_{a}^{(2)}\right) = 2\pi hn, \ n \in Z. \quad (21)$$
Recall that the relative $+$ sign between the two $U(1)$ contributions is due to our identical coupling conventions for both. The only choice that makes (21) consistent with (20) is the $+$ sign as is easy to verify by using the explicit values of the $D = 4$ charges in terms of the $D = 6$ ones. To show, finally, that the electric and magnetic charges of a single string in $D = 6$ are constrained by $2eg = 2\pi nh, \ n \in Z$, we recall that this condition was obtained in [5] by exploiting the flexibility of Dirac membranes to perform motions that do not distinguish between the spatial directions. It comes as no surprise therefore, that one can recover this relation from the $D = 4$ point of view by using $Y$ symmetry. Indeed, together with the configuration $(e,0,0,g)$, $Y$ implies that the configuration $(0,-g,e,0)$ should also exist. Applying the $D = 4$ dyon quantization condition to the two configuration appearing above, we recover this $e - g$ relation. We can imagine continuing this chain of arguments inductively: (1) consider the dyonic configuration in $2p$ dimensions; (2) list the $2p - 2$ dimensional configurations to which it gives rise, including those related by $Y$-symmetry; (3) apply the $2p - 2$ dimensional quantization rules which will therefore relate the $2p$ dimensional parameters, etc.

In retrospect, it is not surprising that one can infer the $D = 6$ quantization condition from that in $D = 4$, together with the extra $Y$-symmetry it enjoys. Indeed, as was shown in [1], the respective quantization conditions with $+/-$ signs possess exactly the same general solutions (assuming existence of pure electric sources); hence (when $(C)P$ invariance is imposed in $D = 4$) they are clearly equivalent [9].

To summarize, we have provided two independent derivations of the $(−)^{p}$ sign factor in the $p$-brane dyon quantization conditions. Both arguments are ultimately manifestations of the basic "double dual" identity $* * = (−)^{p}$.

Acknowledgements

A.S. gratefully acknowledges very stimulating discussions with D. Kutasov; S.D. and M.H. thank A. Gomberoff and C. Teitelboim. The work of S.D. was supported by NSF grant PHY 93-15811, that of M.H. by I.I.S.N. (Belgium), and that of A.S. by the Israeli Academy of Sciences and the Minerva Foun-
dation (Germany). M.H. and A.S. thank Carmen Núñez and the Bariloche High Energy Group for hospitality at the Bariloche Institute, where this work began.

References

An SL(2, Z) multiplet of type IIB super five-branes

J.X. Lu a,1,2, Shibaji Roy b,3

a Center for Theoretical Physics, Physics Department, Texas A & M University, College Station, TX 77843, USA
b Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Calcutta 700 064, India

Received 5 March 1998; revised 19 March 1998
Editor: H. Georgi

Abstract

It is well-known that the low energy string theory admits a non-singular solitonic super five-brane solution which is the magnetic dual to the fundamental string solution. By using the symmetry of the type IIB string theory, we construct an SL(2, Z) multiplet of magnetically charged super five-branes starting from this solitonic solution. These solutions are characterized by two integral three-form charges \( q_1, q_2 \) and are stable when the integers are coprime. We obtain an expression for the tension of these \( (q_1, q_2) \) five-branes as envisaged by Witten. The SL(2, Z) multiplets of black strings and black fivebranes and the existence of similar magnetic dual solutions of strings in type II string theory in \( D < 10 \) have also been discussed.

String theories in the long wavelength limit are described by various kinds of supergravity theories in \( D = 10 \). The supergravity equations of motion are well-known \([1−3]\) to admit various black \( p \)-brane solutions which are essentially the black hole solutions of the dimensionally reduced low energy string effective action spatially extended to the ten dimensional theory. These solutions are usually characterized by two parameters related to the charge and the mass of the black \( p \)-branes. In the extremal limit as usual in the Reissner-Nordstrom black hole, the charge and the ADM mass of the \( p \)-branes are related to each other and the solutions become supersymmetric saturating the BPS bound. The extremal solutions of string effective actions are particularly interesting since their masses and the charges can be calculated exactly due to certain non-renormalization theorems of the underlying supersymmetric theories. Thus although these solutions are obtained from the low energy effective theory, they are quite useful to identify certain non-perturbative symmetries of string theory.

Low energy effective action of any string theory admits a fundamental string solution \([4]\) and its magnetic dual the non-singular solitonic five-brane solution \([5−9]\). These are the extremal limit, as we have mentioned, of the black string and black five-brane solutions of the corresponding supergravity equations of motion. Since it is
known that the equations of motion of type IIB supergravity theory is invariant [10] under an SL(2, R) group one can construct a more general string like as well as five-brane solutions using this symmetry group. We should like to mention that the SL(2, R) symmetry of type IIB string theory is a non-perturbative symmetry. It transforms the string coupling constant in a non-trivial way and therefore mixes up the perturbative and non-perturbative effects of Type IIB string theory. A discrete subgroup of this SL(2, R) group has been conjectured to be the exact symmetry group of the quantum type IIB string theory [11]. Using this symmetry, Schwarz [12] has constructed an SL(2, Z) multiplet of string like solutions in type IIB string theory starting from the fundamental string solution. Both the string tension and the charge were shown to be given by the SL(2, Z) covariant expressions. Since the tension and the charge of these extremal solutions remain unrenormalized, it provides a strong support in favor of the conjecture that SL(2, Z) is an exact symmetry group of the quantum theory.

Given the symmetry of the type IIB theory, it is natural to expect that the solitonic five-brane solution should also form an SL(2, Z) multiplet as pointed out by Schwarz in Ref. [12]. Here we construct an infinite family of magnetically charged super five-branes, permuted by SL(2, Z) group, starting from the solitonic five-brane solution of string theory. These solutions are characterized by a pair of integers corresponding to the magnetic charges associated with the two three-form field strengths present in the NSNS and RR sector of the spectrum. When these two integers \((q_1, q_2)\) are relatively prime to each other, the five-brane solutions are shown to be stable and can be regarded as bound state configuration of \(q_1\) solitonic five-branes with \(q_2\) D5-branes [16]. The magnetic charge as well as the five-brane tension are shown to be given by SL(2, Z) covariant expressions. This provides more evidence that SL(2, Z) is indeed an exact symmetry group of the quantum theory. The expression for the tension of these \((q_1, q_2)\) super five-branes has been envisaged by Witten [15] some time ago. We then discuss that similar magnetic dual solutions of strings also exist in lower dimensional type II theory.

The low energy effective action common to any ten dimensional string theory has the form:

\[
S = \int d^{10}x \sqrt{-g} \left[ R - \frac{1}{2} \nabla_{\mu} \phi \nabla^{\mu} \phi - \frac{1}{12} e^{-\phi} H^{(1)}_{\mu} H^{(1)}_{\mu} \right]
\]  

(1)

Here \(g = (\det g_{\mu\nu})\), \(g_{\mu\nu}\) being the canonical metric which is related to string metric by \(G_{\mu\nu} = e^{\phi/2} g_{\mu\nu}\). \(R\) is the scalar curvature with respect to the canonical metric, \(\phi\) is the dilaton and \(H^{(1)}_{\mu}H^{(1)}_{\mu}\) is the field strength associated with the Kalb-Ramond antisymmetric tensor field \(B^{(1)}_{\mu}\). These are the massless modes which couple to any string theory. For type II strings these massless modes belong to the NSNS sector of the spectrum. The equations of motion following from (1) are:

\[
\nabla_{\mu} (e^{-\phi} H^{(1)}_{\mu}) = 0
\]

(2)

\[
\nabla^2 \phi + \frac{1}{12} e^{-\phi} (H^{(1)})^2 = 0
\]

(3)

\[
R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\frac{1}{4} e^{-\phi} H^{(1)}_{\mu} H^{(1)}_{\nu} + \frac{1}{48} e^{-\phi} (H^{(1)})^2
\]

(4)

These low energy field equations can be solved by using certain ansatz on the metric and the field strength \(H^{(1)}_{\mu}\). By demanding that the metric be static, spherically symmetric which becomes flat asymptotically with a

---

\footnote{Earlier attempt for the construction of SL(2, Z) multiplet of five-brane solution of type IIB string theory were made in [13,14], but the solutions described there are incomplete.}
regular horizon, one can obtain both the electrically charged black string solution and the magnetically charged black five-brane solution from (2)-(4) as given below [3]

\[
\begin{align*}
    ds^2 &= -\left(1 - \frac{r_0^6}{r^6}\right)^{-1/4} dt^2 + \left(1 - \frac{r_0^6}{r^6}\right)^{3/4} \left(dx^i\right)^2 + \left(1 - \frac{r_0^6}{r^6}\right)^{-1} \left(1 - \frac{r_0^6}{r^6}\right)^{-11/12} dr^2 \\
    &+ r^2 \left(1 - \frac{r_0^6}{r^6}\right)^{1/12} d\Omega_7^2 \\
    e^{2\phi} &= \left(1 - \frac{r_0^6}{r^6}\right), \quad H^{(1)} = 6(r_+ r_-)^3 e^\phi \epsilon_7
\end{align*}
\]
for the string solution and

\[
\begin{align*}
    ds^2 &= -\left(1 - \frac{r_0^2}{r^2}\right)^{-3/4} dt^2 + \left(1 - \frac{r_0^2}{r^2}\right)^{1/4} \delta_{ij} dx^i dx^j + \left(1 - \frac{r_0^2}{r^2}\right)^{-1} \left(1 - \frac{r_0^2}{r^2}\right)^{-3/4} dr^2 \\
    &+ r^2 \left(1 - \frac{r_0^2}{r^2}\right)^{1/4} d\Omega_7^2 \\
    e^{-\phi} &= \left(1 - \frac{r_0^2}{r^2}\right), \quad H^{(1)} = 2(r_+ r_-) \epsilon_3
\end{align*}
\]
for the five-brane solution. Here \(i, j = 1, 2, \ldots, 5\). \(d\Omega_7^2\) and \(d\Omega_3^2\) above are the metric on the unit seven dimensional and three dimensional spheres respectively. \(\epsilon_7\) and \(\epsilon_3\) are the corresponding volume forms. The `*` denotes the Hodge duality transformation. \(r_+\) and \(r_-\) are the two parameters related to the mass and the charge of the solutions. Thus Eqs. (5) and (6) represent the two parameter family of black string and black five-brane solutions of string theory with an event horizon at \(r = r_+\) and an inner horizon \(r = r_-\) (where \(r_+ \geq r_-\)). It is clear from the form of \(H^{(1)}\) in (5) that the string solution is electrically charged whereas from (6) we note that the five-brane solution is magnetically charged. We would like to mention that if the supergravity action contains a general \((d + 1)\)-form field strength then the magnetically charged solution can be obtained from electrically charged solution by using the duality transformation \(\phi \rightarrow -\phi, \quad e^{-\phi} \ast H^{(1)} \rightarrow H^{(1)}\) and \(d \rightarrow 8 - d\). Thus the five-brane solution can be obtained from the string solution using this duality transformation. Since, \(e^\phi\) is the string coupling constant, the magnetically charged five-brane solution is a non-perturbative solution of weakly coupled string theory. Note that the solutions (5) and (6) are written assuming that the dilaton vanishes in the asymptotic limit, but we will restore the asymptotically constant value of the dilaton when we write the more general type IIB solution. Finally, we note that for \(r_+ > r_-\), the solutions are non-extremal and therefore non-BPS, but when \(r_+ = r_-\) the solutions become supersymmetric saturating the BPS limit.

In the extremal limit the BPS saturated string solution given in (5) (with \(r_+ = r_-\)), was constructed previously by Dabholkar et. al. [4]. By going to the isotropic coordinate \(\rho^2 = r^6 - r_0^6\), we can rewrite the metric (5) in the extremal limit as,

\[
\begin{align*}
    ds^2 &= \left(1 + \frac{r_0^6}{\rho^6}\right)^{-3/4} \left[(-dt)^2 + (dx^i)^2\right] + \left(1 + \frac{r_0^6}{\rho^6}\right)^{1/4} \left(d\rho^2 + \rho^2 d\Omega_7^2\right)
\end{align*}
\]
with \(e^{-\phi} = \left(1 + \frac{r_0^6}{\rho^6}\right)^{1/2}\). This is precisely the solution discussed in Ref. [4] and clarifies the relation with the solution described in [3]. It is clear from (7) that in terms of the string metric \(G_{\mu\nu} = \left(1 + \frac{\rho^4}{\rho^6}\right)^{-1/4} g_{\mu\nu}\), the above metric becomes flat transverse to the string direction. Also note that the solution (7) is singular since for \(\rho \rightarrow 0\), the radius of \(S^7\) vanishes and the curvature blows up as \(\rho^{-2}\). This is the reason that string like solution has been constructed [4] by coupling the supergravity action to a macroscopic string source. This BPS saturated singular string like solution is also known as the fundamental string solution. By using the symmetry of the full type IIB
string theory including the RR sector, Schwarz has constructed an infinite family of string like solutions starting from this fundamental string solution. In Ref. [17], we have pointed out that a similar infinite family of string solutions also exist in $D < 10$ type IIB string theory.

Let us next look at the black five-brane solution of string theory given in (6). Note that in general the solution is invariant under the symmetry group $R \times E(5) \times SO(4)$ where $E(5)$ is the five dimensional Euclidean group. At the extremal limit, the solution acquires an extra boost symmetry and thus the symmetry group becomes $P(6) \times SO(4)$, where $P(6)$ is the six dimensional Poincare group. So, only at the extremal limit the solution describes the BPS super five-brane and takes the following form:

$$ ds^2 = \left( 1 - \frac{r^2}{r_s^2} \right)^{1/4} \left[ -(dt)^2 + \delta_{ij} dx^i dx^j \right] + \left( 1 - \frac{r^2}{r_s^2} \right)^{-7/4} dr^2 + r^2 \left( 1 - \frac{r^2}{r_s^2} \right)^{1/4} d\Omega_3^2 $$

$$ Q = 2r_s^2, \quad e^{-2\phi} = \left( 1 - \frac{r^2}{2r_s^2} \right), \quad H^{(3)} = Qe_3 \tag{8} $$

In terms of the isotropic coordinate $\rho^2 = r^2 - r_s^2$ the solution can be written as,

$$ ds^2 = \left( 1 + \frac{r^2}{\rho^2} \right)^{-1/4} \left[ -(dt)^2 + \delta_{ij} dx^i dx^j \right] + \left( 1 + \frac{r^2}{\rho^2} \right)^{3/4} \left( dp^2 + \rho^2 d\Omega_3^2 \right) $$

$$ Q = 2r_s^2, \quad e^{2\phi} = \left( 1 + \frac{Q}{2r_s^2} \right)^{1/4}, \quad H^{(1)} = Qe_3 \tag{9} $$

In the string frame $G_{\mu\nu} = \left( 1 + \frac{r^2}{\rho^2} \right)^{1/4} g_{\mu\nu}$, the metric in (9) reduces to

$$ ds^2 = \left[ -(dt)^2 + \delta_{ij} dx^i dx^j \right] + \left( 1 + \frac{r^2}{\rho^2} \right) \left( dp^2 + \rho^2 d\Omega_3^2 \right) \tag{10} $$

Note that unlike the string solution, the super five-brane solution is regular as $\rho \to 0$ since in this case the radius of $S^3$ is finite ($r_s$) and the curvature also remains finite ($r_s^{-2}$) as $\rho \to 0$.

We would like to point out that the low energy string effective action we have considered in (1) can be regarded as a special case of more general type IIB action when the RR fields are included. Let us recall that the massless states of type IIB string theory in the bosonic sector consist of a graviton ($G_{\mu\nu}$), a dilaton ($\phi$) and an antisymmetric tensor field ($B^{(1)}_{\mu\nu}$) as NSNS gauge fields whereas in the RR sector it has another scalar ($\chi$), another antisymmetric tensor field ($B^{(2)}_{\mu\nu}$) and a four-form gauge field ($A^{+}_{\mu\nu\rho\sigma}$) whose field strength is self-dual. It is well-known that the equations of motion of type IIB supergravity theory is invariant under an $SL(2, R)$ group known as the supergravity duality group [10]. The four-form gauge field is a singlet under this duality group and it couples to a self-dual three-brane whose form has been derived in Ref. [3,18]. Since we are interested in the five-brane solution we set the corresponding five-form field strength to zero in what follows. In this case the type IIB supergravity equations of motion can be derived from the following covariant action [19],

$$ S_{\text{IIB}} = \int d^{10}x \sqrt{-G} \left[ e^{-2\phi} \left( R + 4 \nabla_{\mu} \phi \nabla^{\mu} \phi - \frac{1}{12} H^{(1)}_{\mu\nu\lambda} H^{(1)}_{\mu\nu\lambda} - \frac{1}{2} \nabla_{\mu} \chi \nabla^{\mu} \chi \right) + \frac{1}{2} H^{(2)}_{\mu\nu\lambda} H^{(2)}_{\mu\nu\lambda} + \chi H^{(1)}_{\mu\nu\lambda} \right] \tag{11} $$

We can rewrite the action (11) in the Einstein frame as follows:

$$ S_{\text{IIB}} = \int d^{10}x \sqrt{-g} \left[ R - \frac{1}{2} \nabla_{\mu} \phi \nabla^{\mu} \phi - \frac{1}{2} e^{2\phi} \nabla_{\mu} \chi \nabla^{\mu} \chi \right. $$

$$ \left. - \frac{1}{12} \left( e^{-\phi} H^{(1)}_{\mu\nu\lambda} H^{(1)}_{\mu\nu\lambda} + e^{\phi} (H^{(2)}_{\mu\nu\lambda} + \chi H^{(1)}_{\mu\nu\lambda})(H^{(2)}_{\mu\nu\lambda} + \chi H^{(1)}_{\mu\nu\lambda}) \right) \right] \tag{12} $$
It is to be noted that (12) reduces precisely to the effective action (1) when the RR fields are set to zero. Since (1) is a special case of (12), the five-brane solution obtained from (1) (given in (6), (8) or (9)) can also be generalized for the type IIB theory. We are going to construct in the following these generalized solution of five-branes of type IIB theory. The construction will be facilitated if we write the action (12) in the manifestly SL(2, R) invariant form as given below [12,19]:

\[ S_{\text{IB}} = \int d^{10} x \sqrt{-g} \left[ R + \frac{1}{4} \text{tr} \nabla_{\mu} \mathcal{M} \nabla^{\mu} \mathcal{M}^{-1} - \frac{1}{12} \mathcal{M}^{\alpha \beta} \mathcal{M}^{\mu \nu \lambda} \mathcal{M}^{\mu \nu \lambda} \right] \] (13)

where \( \mathcal{M} \equiv \begin{pmatrix} \chi^2 + e^{-2 \phi} & \chi \\ \chi & 1 \end{pmatrix} \) represents an SL(2, R) matrix and \( \mathcal{M}^{\mu \nu \lambda} \equiv \begin{pmatrix} H_{\mu \nu}^{(1)} \\ H_{\mu \nu}^{(2)} \end{pmatrix} \). Also the superscript ‘\( T \)’ represents the transpose of a matrix. The action (13) can be easily seen to be invariant under the following global SL(2, R) transformations:

\[ \mathcal{M} \rightarrow A \mathcal{M} A^T, \quad \mathcal{M}_{\mu \nu} \rightarrow (A^{-1})^T \mathcal{M}_{\mu \nu}, \quad g_{\mu \nu} \rightarrow g_{\mu \nu} \] (14)

where \( A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \), with \( ad - bc = 1 \), represents a global SL(2, R) transformation matrix. It is easy to check that under the transformation (14), the complex scalar \( \lambda = \chi^2 + e^{-2 \phi} \) and the two field strengths \( H_{\mu \nu}^{(1)} \) and \( H_{\mu \nu}^{(2)} \) transform as,

\[ \lambda \rightarrow \frac{a \lambda + b}{c \lambda + d}, \quad H_{\mu \nu}^{(1)} \rightarrow dH_{\mu \nu}^{(1)} - cH_{\mu \nu}^{(2)}, \quad H_{\mu \nu}^{(2)} \rightarrow -bH_{\mu \nu}^{(1)} + aH_{\mu \nu}^{(2)} \] (15)

We would like to point out here that unlike the electrically charged string solution, the magnetic charges associated with \( H_{\mu \nu}^{(1)} \) and \( H_{\mu \nu}^{(2)} \) of the five-brane should transform in the same way as the field strengths themselves. This follows from the fact that Noether charge (or the electric charge) of the string solution is conserved due to the equation of motion following from (13) whereas the topological charge (or the magnetic charge) of the five-brane is conserved due to Bianchi identity. Therefore the magnetic charges of the five-branes transform as \( Q \rightarrow (A^{-1})^T Q \) or in components,

\[ Q^{(1)} \rightarrow dQ^{(1)} - cQ^{(2)}, \quad Q^{(2)} \rightarrow -bQ^{(1)} + aQ^{(2)} \] (16)

Note that the original solution (6) or (9) had one charge \( Q \) associated with \( H^{(1)} = Qe_x \), and this charge was quantized in some basic units. Now after the transformation (16) \( Q \) will no longer remain quantized. So, in order to recover the charge quantization [20] we modify the original charge by \( \Delta_{(q_1, q_2)}^{1/2} Q \), where \( \Delta_{(q_1, q_2)}^{1/2} \) is an arbitrary constant which will be determined later. The construction of the SL(2, R) matrix \( A \) can be motivated such that it properly reproduces the asymptotic value of the complex scalar \( \lambda_0 = \chi_0^2 + e^{-2 \phi_0} \), after the transformation, where \( \phi_0 \) and \( \chi_0 \) are the asymptotic value of the dilaton and the RR scalar. The relevant SL(2, R) transformation matrix then takes the form [12,17]

\[ A = \begin{pmatrix} e^{-\phi_0 / 2} \cos \alpha + \chi_0 \sin \alpha & -e^{-\phi_0 / 2} \sin \alpha + \chi_0 \cos \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} e^{\phi_0 / 2} \] (17)

Here \( \alpha \) is an arbitrary parameter which will be fixed from the charge quantization condition. From (16), we find the charges associated with \( H_{\mu \nu}^{(1)} \) and \( H_{\mu \nu}^{(2)} \) to be given as,

\[ Q^{(1)} = e^{\phi_0 / 2} \cos \alpha \Delta_{(q_1, q_2)}^{1/2} Q, \quad Q^{(2)} = (e^{-\phi_0 / 2} \sin \alpha - \chi_0 e^{\phi_0 / 2} \cos \alpha) \Delta_{(q_1, q_2)}^{1/2} Q \] (18)

By demanding that the charges be quantized we find,

\[ \cos \alpha = e^{-\phi_0 / 2} \Delta_{(q_1, q_2)}^{-1/2} q_1, \quad \sin \alpha = e^{\phi_0 / 2} (q_2 + q_1 \chi_0) \Delta_{(q_1, q_2)}^{-1/2} \] (19)
where \( q_1 \) and \( q_2 \) are integers. Using \( \cos^2 \alpha + \sin^2 \alpha = 1 \), (19) determines the value of \( \Delta_{(q_1, q_2)} \) as,

\[
\Delta_{(q_1, q_2)} = e^{-\phi_0} q_1^2 + (q_2 + q_1 \chi_0)^2 e^{\phi_0} = (q_1, q_2) \mathcal{M}_0 \left( \frac{q_1}{q_2} \right)
\]  

(20)

where \( \mathcal{M}_0 = \left( \frac{\chi_0^2 + e^{-2\phi_0}}{\chi_0} \right) e^{\phi_0} \). It is clear from (20) that \( \Delta_{(q_1, q_2)} \) is SL(2, \( \mathbb{Z} \)) covariant. Therefore, the charge of a \((q_1, q_2)\) five-brane is proportional to \( \frac{1}{\chi_0} \). We can also calculate the tension of a \((q_1, q_2)\) five-brane by an SL(2, \( \mathbb{Z} \)) covariant expression

\[
Q_{(q_1, q_2)} = \Delta_{(q_1, q_2)}^{1/2} Q = \sqrt{e^{-\phi_0} q_1^2 + (q_2 + q_1 \chi_0)^2 e^{\phi_0}} \quad Q
\]  

(21)

Note from (14) that the canonical metric does not change under the SL(2, \( \mathbb{R} \)) transformation. However, since the charge \( Q \) is now replaced by \( Q_{(q_1, q_2)} \) the metric given in (9) takes the following form:

\[
ds^2 = \left( 1 + \frac{Q_{(q_1, q_2)}}{2\rho^2} \right)^{1/4} \left[ -(dt)^2 + \delta_{ij} dx^i dx^j \right] + \left( 1 + \frac{Q_{(q_1, q_2)}}{2\rho^2} \right)^{3/4} \left( d\rho^2 + \rho^2 d\Omega_5^2 \right)
\]  

(22)

The complex scalar field \( \lambda \) changes as

\[
\lambda = \frac{a(e^{-\phi_0}) + b}{c(e^{-\phi_0}) + d} = \frac{\chi_0 \Delta_{(q_1, q_2)} A_{(q_1, q_2)} + q_1 q_2 e^{-\phi_0}(A_{(q_1, q_2)} - 1) + i \Delta_{(q_1, q_2)} A_{(q_1, q_2)} e^{-\phi_0}}{q_1^2 e^{-\phi_0} + A_{(q_1, q_2)} e^{\phi_0}(\chi_0 q_1 + q_2)^2}
\]  

(23)

where \( A_{(q_1, q_2)} = \left( 1 + \frac{Q_{(q_1, q_2)}}{2\rho^2} \right)^{-1} \). Note that asymptotically as \( \rho \to \infty \), \( A_{(q_1, q_2)} \to 1 \) and therefore, \( \lambda \to \lambda_0 \) as expected. The real and imaginary part of (23) give the transformed value of the RR scalar and the dilaton of the theory. Finally, the transformed value of the field strengths \( H^{(1)} \) and \( H^{(2)} \) can be obtained from (15) as,

\[
H^{(1)} = q_1 Q e_1, \quad H^{(2)} = q_2 Q e_3
\]  

(24)

which can be written compactly as follows,

\[
\mathcal{F} = \left( \frac{q_1}{q_2} \right) \lambda Q e_3
\]  

(25)

We can also calculate the tension of a \((q_1, q_2)\) five-brane by calculating the ADM mass per unit five-volume [21]. We note that in general the ADM mass is given by \( M = (3r_+^2 - r_-^2) \) and therefore, for the non-extremal case the mass and the charge are independent parameters. But in the extremal case they are related since in that case, \( M = 2r_-^2 = Q \). We have seen in (21) that the charge of a \((q_1, q_2)\) five-brane has been modified by \( \Delta_{(q_1, q_2)}^{1/2} \) \( Q \) and so in order to equate the mass with the charge, mass per unit five-volume i.e. the tension must also satisfy the similar relation:

\[
T_{(q_1, q_2)} = \Delta_{(q_1, q_2)}^{1/2} T = \sqrt{e^{-\phi_0} q_1^2 + (q_2 + q_1 \chi_0)^2 e^{\phi_0}} T = \sqrt{g_s^{-1} q_1^2 + (q_2 + q_1 \chi_0)^2} \ T
\]  

(26)

where \( g_s = e^{\phi_0} \) in the last expression of (26) denotes the string coupling constant. Thus when \( \chi = 0 \), the solitonic five-brane or \((1, 0)\) brane tension is proportional to \( 1/\sqrt{g_s} \) whereas D5-brane or \((0, 1)\) brane tension is proportional to \( \sqrt{g_s} \) in the canonical metric. In the string metric, on the other hand, the tension of a general \((q_1, q_2)\) five-brane is given by

\[
T_{(q_1, q_2)} = g_s^{-3/2} \sqrt{g_s^{-1} q_1^2 + g_s q_2^2} \ T
\]  

(27)

Here \( T_{(q_1, q_2)} \) gets scaled by \( g_s^{-3/2} \) because, it has the dimensionality of \((\text{length})^{-3}\). Thus in the string metric, the tension of a solitonic five-brane is proportional to \( 1/g_s^2 \) and the tension of a D5-brane is proportional to \( 1/g_s \) as expected. This tension formula of a \((q_1, q_2)\) five-brane has been envisaged by Witten in Ref. [15].
Thus starting from the solitonic five-brane solution, we have obtained an infinite family of five-brane solutions permitted by SL(2, Z) group in type IIB theory given by the metric and other field configurations in (22) – (25).

The stability of these \((q_1,q_2)\) five-brane solutions can be understood along the same line as in the case of string solutions [17,21,22]. Since the tension of a \((q_1,q_2)\) five-brane is given in (26), it can be easily checked that when \(\chi = 0\), the five-brane tension satisfy the following triangle inequality

\[
T_{(p_1,q_1)} + T_{(q_1,q_2)} \geq T_{(p_1+q_1,q_2+q_3)}
\]

(28)

Such relation is quite typical of a BPS state. The equality holds when \(p_1q_2 = p_2q_1\) or in other words when \(p_1 = nq_1\) and \(p_2 = nq_2\), where \(n\) is any integer. Thus when \(q_1\) and \(q_2\) are relatively prime, the inequality prevents the five-brane state to decay into five-branes of lower masses. Since the charge of a \((q_1,q_2)\) five-brane also satisfies similar relation (21) it can be readily checked again that when \(q_1\), \(q_2\) are coprime the charge conservation can not be satisfied if the five-brane decay into multiple five-branes. Thus \((q_1,q_2)\) five-brane configuration with \(q_1\), \(q_2\) relatively prime, describes a bound state configuration of \(q_1\) solitonic five-branes with \(q_2\) D5-branes.

Note here, as discussed by Witten, that unlike the string solution, the D5-branes themselves do not form bound states as the six dimensional super Yang-Mills theory does not contain vacua with a mass gap. On the other hand, D5-branes when combined with solitonic five-branes do form bound states and has been discussed qualitatively by Witten in Ref. [15].

The SL(2, Z) multiplet of black fivebranes can be obtained similarly. Eqs. (14)–(21) and Eqs. (23)–(25) remain the same but (22) should be replaced by (6) with \(r_+\) and \(r_-\) now given by \(2(r_-,r_+) = \Delta r_{(q_1,q_2)}\). The same procedure applies to the construction of the SL(2, Z) multiplet of black strings. These multiplets may be useful in studying the physics of black strings and fivebranes.

Finally, we would like to mention that similar infinite family of magnetic dual solutions of strings also exist in \(D > 4\). The low energy effective action common to any string theory in \(D\) dimensions has the form:

\[
S_\Omega = \int d^Dx \sqrt{-g} \left[ R - \frac{4}{D-2} \nabla^\mu \phi \nabla_\mu \phi - \frac{1}{12} e^{- \frac{8}{D-2} \phi} H_{\mu \nu \lambda}^4 H^{(I) \mu \nu \lambda} \right]
\]

(29)

The equations of motion following from (29) are given as:

\[
\nabla^\mu \left( e^{- \frac{8}{D-2} \phi} H^{(I) \mu \nu \lambda} \right) = 0
\]

(30)

\[
\nabla^2 \phi + \frac{1}{12} e^{- \frac{8}{D-2} \phi} (H^{(I)})^2 = 0
\]

(31)

\[
R_{\mu \nu} - \frac{4}{D-2} \nabla_\mu \phi \nabla_\nu \phi - \frac{1}{4} e^{- \frac{8}{D-2} \phi} H_{\mu \rho \sigma}^4 H^{(I) \rho \sigma} + \frac{1}{6(D-2)} e^{- \frac{8}{D-2} \phi} (H^{(I)})^2 g_{\mu \nu} = 0
\]

(32)

By using the same ansatz on the metric as before one can obtain the magnetic dual solution of the string in \(D > 4\) dimensions of the form:

\[
ds_\Omega^2 = A^{-\frac{2}{D-2}} \left[ -(dt)^2 + \delta_{ij} (dx^i dx^j) \right] + A^{\frac{D-4}{2}} (d\rho^2 + \rho^2 d\Omega_7^2)
\]

(33)

where \(A\) is a function of radial coordinate \(\rho\) only whose explicit form is given in [9] as:

\[
e^{2\phi} = (A(\rho))^{\frac{D-2}{8}} \left( 1 + \frac{r_-^2}{\rho^2} \right)^{\frac{D-2}{8}}
\]

(34)

where \(r_-\) is the charge of the dual object. Also, the field strength \(H^{(I)} = Q\epsilon_5\). The magnetic dual object is a 4-brane in \(D = 9\), a 3-brane in \(D = 8\), a 2-brane in \(D = 7\), a string in \(D = 6\) and a 0-brane (a particle) in \(D = 5\).
Since it is known that the toroidally compactified type IIB string theory also possesses the $SL(2, \mathbb{R})$ invariance (23,24) This symmetry can also be obtained in $D \leq 9$ from toroidally compactified M-theory, see, e.g., [25]. With the same transformation properties of $\mathcal{F}_{\mu\nu}$ and the complex scalar $\lambda$ and since the reduced action can be shown to be given by (29) when RR fields are set to zero, we can straightforwardly use the $SL(2, \mathbb{Z})$ rotation to find the $SL(2, \mathbb{Z})$ family of the dual solutions of strings starting from (33) and (34). The solutions in this case are very similar as in the ten dimensional case. (For the detailed construction of string solution in $D < 10$ see [17].) The corresponding black $SL(2, \mathbb{Z})$ multiplet for each of the dual objects can be constructed in a similar fashion described above.

To conclude, we have constructed in this paper an $SL(2, \mathbb{Z})$ family of super five-brane solutions in type IIB string theory starting from the known non-singular solitonic five-brane solution. These solutions are characterized by a pair of integers corresponding to the magnetic charges associated with the two three-form field strengths in the NSNS and RR sector of the theory. We have shown that both the charge and the tension of a general $(q_1, q_2)$ five-brane are given by $SL(2, \mathbb{Z})$ covariant expressions. This provides more evidence in support of the conjecture that $SL(2, \mathbb{Z})$ is an exact symmetry group of the quantum type IIB string theory. When the integers $q_1$ and $q_2$ are relatively prime, we have shown that the five-brane is stable since it is prevented from decaying into multiple five-branes by a triangle inequality relation of both the tension as well as the charge associated with the five-brane. We have obtained the tension formula for a general $(q_1, q_2)$ five-brane as envisaged by Witten in Ref. [15]. We have discussed that a similar family of the magnetic-dual solutions of the string also exists in each of $D > 4$ dimensions in type II string theory. We also discussed how to obtain the corresponding $SL(2, \mathbb{Z})$ multiplets for black strings and its dual objects for $10 \geq D > 4$.

References

On generalized axion reductions

Nemanja Kaloper a,1, Ramzi R. Khuri b,2, Robert C. Myers c,3,4

1 Department of Physics, Stanford University, Stanford, CA 94305-4060, USA
2 Department of Physics, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK
3 Institute of Theoretical Physics, University of California, Santa Barbara, CA 93106-4030, USA

Received 9 March 1998
Editor: M. Cvetic

Abstract

Recently interest in using generalized reductions to construct massive supergravity theories has been revived in the context of M-theory and superstring theory. These compactifications produce mass parameters by introducing a linear dependence on internal coordinates in various axionic fields. Here we point out that by extending the form of this simple ansatz, it is always possible to introduce the various mass parameters simultaneously. This suggests that the various “distinct” massive supergravities in the literature should all be a part of a single massive theory. © 1998 Elsevier Science B.V. All rights reserved.

Dimensional reduction provides an important window on the duality relations amongst the various superstring theories, as well as eleven-dimensional supergravity. Recently generalized Scherk-Schwarz reductions [1] have received a renewed interest [2–6]. This activity began with the remarkable discovery [2] that the massive IIA supergravity of Romans [7] is related by T-duality to a Scherk-Schwarz compactification of the massless IIB theory. This result then provides a massive extension of the standard T-duality between type IIA and IIB superstring theories compactified on $S^1$ [8]. Further, renewed interest stems from the recent investigations of extended objects in string theory. Massive supergravities are particularly relevant in the case of domain walls [3,4]. Some earlier investigations of Scherk-Schwarz reductions in string theory were made both at the level of the low-energy supergravity action [9], and at the level of the world-sheet conformal field theory [10].

The key to the generalized Scherk-Schwarz reductions [1,6] is that, using global symmetries arising in a compactification, the fields may be given a (specific) dependence on the internal coordinates. However, the resulting theory is still independent of all of the internal coordinates. The recent discussions [2–5] in the context of low-energy string or M-theory focus on toroidal compactifications and various axionic symmetries, i.e., constant shifts of certain scalar fields. In the simplest cases then, the axions appear in the action covered by

---

1 E-mail: kaloper@leland.stanford.edu.
2 E-mail: R.R.Khuri@qmw.ac.uk.
3 E-mail: rcm@itp.ucsb.edu.
4 Permanent address: Department of Physics, McGill University, Montréal, PQ, H3A 2T8, Canada.

0370-2693/98/$19.00 © 1998 Elsevier Science B.V. All rights reserved.
PII: S0370-2693(98)00409-2
derivatives, i.e., the scalar field $\chi$ appears everywhere in the action only as $\partial_\mu \chi$, or in form notation as $d\chi$. If upon compactification such axions are given a linear dependence on the internal coordinates, only the slope of this dependence appears in the reduced action [3], i.e.,

$$\chi(x,z) = \chi(x) + mz \rightarrow d\chi(x,z) = d\chi(x) + m dz. \quad (1)$$

The slope parameters then play the role of masses in the compactified theory.

A fundamental axion scalar appears in the ten-dimensional IIb supergravity and plays the central role in the T-duality to the massive IIa theory [2]. In general, however, the axions of interest arise in a partially reduced theory as internal components of gauge fields, form-fields or the metric. The translation symmetry of these scalars is then a residue of a local gauge invariance in the uncompactified theory. Introducing the linear ansatz $\chi$ to several axions, but were limited by the problem discussed above. The present discussion provides an explicit extension of their results, and generalizing this approach to other cases is a straightforward exercise.

Cowdall et al. [3] applied Scherk-Schwarz reductions to eleven-dimensional supergravity to produce a variety of maximally-supersymmetric massive supergravities in $D \leq 8$. They discussed the case of simultaneously applying the linear ansatz (1) to several axions, but were limited by the problem discussed above. The present discussion provides an explicit extension of their results, and generalizing this approach to other cases is a straightforward exercise.

In this letter, we make the discussion explicit by referring to a specific example considered in Ref. [3]. Cowdall et al. [3] applied Scherk-Schwarz reductions to eleven-dimensional supergravity to produce a variety of maximally-supersymmetric massive supergravities in $D \leq 8$. They discussed the case of simultaneously applying the linear ansatz (1) to several axions, but were limited by the problem discussed above. The present discussion provides an explicit extension of their results, and generalizing this approach to other cases is a straightforward exercise.

In the toroidal compactification of eleven-dimensional supergravity to $D = 8$, three axions $\mathcal{A}_0^{(i)}$ (with $i, j = 1, 2, 3$ and $i < j$) appear in the off-diagonal components of the internal metric. The appropriate dreibein on the internal torus may be written (using the notation of [11])

$$e^A_M = \begin{pmatrix} e^{-\phi_1} & e^{-\phi_1} & e^{-\phi_1} & e^{-\phi_1} \\ 0 & e^{-\phi_2} & e^{-\phi_2} & e^{-\phi_2} \\ 0 & 0 & e^{-\phi_3} & e^{-\phi_3} \end{pmatrix}, \quad (2)$$

where $A$ and $M$ denote the tangent-space and holonomic indices, respectively. The kinetic terms of the axions are governed by the "field strengths"

$$\mathcal{F}^{(12)} = d\mathcal{A}_0^{(12)} \quad \mathcal{F}^{(13)} = d\mathcal{A}_0^{(13)} - d\mathcal{A}_0^{(23)}d\mathcal{A}_0^{(12)} \quad \mathcal{F}^{(23)} = d\mathcal{A}_0^{(23)}. \quad (3)$$

It is clear here that upon compactifying to $D = 7$ one can straightforwardly introduce the linear ansatz (1) for $\mathcal{A}_0^{(12)}$ and $\mathcal{A}_0^{(13)}$. To apply this ansatz to $\mathcal{A}_0^{(23)}$, one must redefine the fields [3] as $\mathcal{A}_0^{(13)} = \mathcal{A}_0^{(13)} - \mathcal{A}_0^{(23)}\mathcal{A}_0^{(12)}$, such that

$$\mathcal{F}^{(13)} = d\mathcal{A}_0^{(13)} + \mathcal{A}_0^{(12)}d\mathcal{A}_0^{(23)}. \quad (4)$$

Now in reducing to $D = 7$, one can apply the ansatz

$$\mathcal{A}_0^{(23)}(x,z) = \mathcal{A}_0^{(23)}(x) + m^{(23)}z. \quad (5)$$

However, from Eq. (4), one sees that this ansatz may no longer be applied to $\mathcal{A}_0^{(12)}$.

\[\text{We have simplified this notation with respect to the scalars } \phi_i, \text{ which do not play an important role in the following.}\]
As an alternative to making the above field redefinition, one could extend the compactification ansatz slightly as follows:

\[
\mathcal{A}_0^{(12)}(x, z) = \mathcal{A}_0^{(12)}(x), \quad \mathcal{A}_0^{(23)}(x, z) = \mathcal{A}_0^{(23)}(x) + m^{(23)}z, \\
\mathcal{A}_0^{(13)}(x, z) = \mathcal{A}_0^{(13)}(x) + m^{(23)}z \mathcal{A}_0^{(12)}(x).
\]

(6)

The additional term added to \( \mathcal{A}_0^{(13)} \) is a reflection of the fact that the axion shift symmetry of \( \mathcal{A}_0^{(23)} \) in the original theory is accompanied by a compensating shift of \( \mathcal{A}_0^{(13)} \) so as to leave \( \mathcal{F}_1^{(13)} \) invariant. We see by replacing (6) into the original expression for the field strengths (3) that all of the explicit \( z \) dependence cancels.

(Alternatively, we note that this ansatz is identical to the original one in which implicitly we have reduced the extended ansatz to ensure that the corresponding field strengths do not introduce a compensating shift of \( \mathcal{F}_1^{(13)} \) in the original theory is accompanied by a compensating shift of \( \mathcal{A}_0^{(13)} \) so as to leave \( \mathcal{F}_1^{(13)} \) invariant. We see by replacing (6) into the original expression for the field strengths (3) that all of the explicit \( z \) dependence cancels.

Within the Scherk-Schwarz formalism, one begins by identifying the relevant global symmetries. Here, they are a part of the SL(3, R) symmetry acting on the internal three-torus, which acts on the dreibein (2) as \( e^A_M \rightarrow e^A_N T^N_M \). The translations of the axions can be identified as the three transformations with generators

\[
M^{(12)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M^{(13)} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M^{(23)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\]

(7)

For example, \( e^A_M \rightarrow e^A_N \exp(\lambda M^{(12)} N_M) \) accomplishes a shift \( \mathcal{A}_0^{(12)} \rightarrow \mathcal{A}_0^{(12)} + \lambda \). Note that \( M^{(23)} \) produces \( \mathcal{A}_0^{(23)} \rightarrow \mathcal{A}_0^{(23)} + \lambda \), and as well, the compensating shift \( \mathcal{A}_0^{(13)} \rightarrow \mathcal{A}_0^{(13)} + \lambda \mathcal{A}_0^{(12)} \). A distinguishing property of these three generators (7) is that they are nilpotent.

Now in the Scherk-Schwarz reduction [1] to \( D = 7 \), one introduces the following specific dependence on the new internal coordinate \( z \) into the dreibein (2)

\[
e^A_N(x, z) = e^A_N(x) U(z)_N^M = e^A_N(x) \exp[M z]_N^M,
\]

(8)

where \( M = \sum m^{(i)} M^{(i)} \). If we consider only a single nonvanishing mass parameter at a time, it is clear that this ansatz reproduces the usual linear ansatz discussed above because the individual generators are nilpotent, i.e., the exponential reduces to \( 1 + M z \). However, in the case that \( m^{(12)} \) and \( m^{(23)} \) are simultaneously chosen to be nonvanishing, \( M^2 = m^{(12)} m^{(23)} M^{(13)} \neq 0 \) while \( M \neq 0 \). Thus in this situation the linear ansatz is naturally extended to one quadratic in the internal coordinate \( z \). Explicitly, the axes are chosen as

\[
\mathcal{A}_0^{(12)}(x, z) = \mathcal{A}_0^{(12)}(x) + m^{(12)}z, \quad \mathcal{A}_0^{(23)}(x, z) = \mathcal{A}_0^{(23)}(x) + m^{(23)}z, \\
\mathcal{A}_0^{(13)}(x, z) = \mathcal{A}_0^{(13)}(x) + m^{(13)}z + m^{(12)}m^{(23)}z^2.
\]

(9)

One can verify that there is no explicit \( z \) dependence in \( \mathcal{F}_1^{(13)} \) with this ansatz. Thus within the full Scherk-Schwarz framework [1], one finds that there is no obstacle to turning on all of the mass parameters simultaneously.

These axes also couple to other fields in the \( D = 8 \) supergravity, and one must also choose a consistent compactification ansatz to ensure that the corresponding field strengths do not introduce a \( z \) dependence in the compactified theory. The Scherk-Schwarz formalism provides a precise prescription to accomplish this result. Essentially any of the fields carrying internal holonomic indices are also contracted with the same matrix \( U \) appearing in Eq. (8). In the present case of the compactification of eleven-dimensional supergravity, one must consider the components of the three-form potential, e.g., \( A_{M N M}(x) U(z)_N^P U(z)_P^M \) and \( A_{M N M}(x) U(z)_N^P \).
Following the notation of [11], these correspond to the one-forms $A_1^{(i)}$ and two-forms $A_1^{(i)}$. In the end, one arrives at the following reduction ansatz for the one-forms:

$$A_1^{(12)}(x,z) = \begin{vmatrix} A_1^{(1)}(x) \\ A_1^{(14)}(x) \end{vmatrix}, \quad A_1^{(13)}(x,z) = \begin{vmatrix} A_1^{(1)}(x) + m^{(12)}zA_1^{(12)}(x) \\ A_1^{(14)}(x) + m^{(23)}zA_1^{(23)}(x) \end{vmatrix},$$

$$A_1^{(23)}(x,z) = \begin{vmatrix} A_1^{(2)}(x) + m^{(12)}zA_1^{(12)}(x) - (m^{(13)}z - \frac{1}{2}m^{(12)}m^{(23)}z^2)A_1^{(23)}(x) \\ A_1^{(24)}(x) + m^{(12)}zA_1^{(24)}(x) - (m^{(13)}z - \frac{1}{2}m^{(12)}m^{(23)}z^2)A_1^{(24)}(x) \end{vmatrix},$$

and for the two-forms:

$$A_2^{(3)}(x,z) = \begin{vmatrix} A_2^{(3)}(x) + m^{(12)}zA_2^{(12)}(x) + (m^{(13)}z + \frac{1}{2}m^{(12)}m^{(23)}z^2)A_2^{(13)}(x) \\ A_2^{(34)}(x) + m^{(12)}zA_2^{(34)}(x) + (m^{(13)}z + \frac{1}{2}m^{(12)}m^{(23)}z^2)A_2^{(34)}(x) \end{vmatrix}.$$  \hspace{1cm} (10)

Here, we see that the quadratic terms make their presence felt in $A_1^{(23)}$ and $A_2^{(3)}$. Again one may explicitly verify that with this ansatz no dependence on $z$ appears in the corresponding field strengths. One must also consider the axion $A_1^{(23)}$ which corresponds to the three-form potential component with three internal indices. Following the Scherk-Schwarz prescription, the compactification ansatz is

$$A_{N_i,N_j,M}(x) U(z)y_i U(z)y_j U(z)y_k = A_{M_i,M_j,M}(x) \det U.$$ \hspace{1cm} (12)

However, $\det U = 1$, so this scalar is unaffected by the above Scherk-Schwarz ansatz. In more general settings, one could not expect such a cancellation to occur. Further, one might also consider the spacetime vectors arising from the off-diagonal components of the eleven-dimensional metric. However, with the present notation of Ref. [11], one does not introduce any $z$ dependence for these vectors – note that the notations of Refs. [11] and [1] differ for these fields.

It should be clear at this point that if we were to extend this discussion to generalized Scherk-Schwarz compactifications to lower dimensions, the linear axion ansatz would again be extended to include cubic and higher order terms in the internal coordinate. We also note, however, that there does remain the possibility of using field redefinitions to simplify the ansatz to one with only linear dependence on the internal coordinates. In the present example, redefining $\omega_0^{(13)} = \omega_0^{(13)} - \frac{1}{z}\omega_0^{(12)}\omega_0^{(23)}$ eliminates the quadratic terms in the compactification ansatz (9) leaving

$$\omega_0^{(13)}(x,z) = \omega_0^{(13)}(x) + m^{(12)}z + \frac{1}{2}m^{(12)}m^{(23)}z^2 \omega_0^{(12)}(x) - \frac{1}{2}m^{(12)}m^{(23)}z^2 \omega_0^{(23)}(x).$$ \hspace{1cm} (13)

Similarly redefining

$$\omega_1^{(23)} = A_1^{(23)} - \frac{1}{2}\omega_0^{(12)}\omega_0^{(23)} A_1^{(12)}, \quad A_2^{(3)} = A_2^{(3)} - \frac{1}{2}\omega_0^{(12)}\omega_0^{(23)} A_2^{(1)},$$ \hspace{1cm} (14)

removes the quadratic terms from Eqs. (11) and (12). Although linear in $z$, this reduction ansatz still does not take the original simple form of Eq. (1).

In summary, one finds that there is no obstacle to simultaneously applying a Scherk-Schwarz reduction for all four of the eight-dimensional axions, $\omega_0^{(12)}$, $\omega_0^{(13)}$, $\omega_0^{(23)}$ and $A_0^{(123)}$ – we have not considered the latter above, but there is no conflict in introducing $m^{(123)}$ along with any of the other mass parameters [3].

\footnote{These field strengths are explicitly listed in Ref. [3]. Note that it is important to explicitly retain certain higher order terms, i.e., $F_i^{(1)} = dA_1^{(1)} - (\omega_0^{(13)} - \omega_0^{(12)}\omega_0^{(23)})dA_1^{(1)} - \omega_0^{(23)}dA_2^{(1)} + \ldots$.}
conclusion also applies to other compactifications. Hence, the various massive supergravities presented in Ref. [3] as distinct theories should actually be regarded as belonging to a single family of theories. After suitable field redefinitions one finds in the present example that generically there is a three-parameter scalar potential involving $m^{123}$, $m^{12}$ and $m^{23}$, while $m^{13}$ can be completely removed from the action. The latter is essentially accomplished by absorbing $m^{13}$ in the expectation value of the axion $\phi_0^{23}$ (as long as $m^{12}$ is nonvanishing) [3]. Given the Scherk-Schwarz framework, one should be able to extend this theory further by beginning with the massive type IIA theory in ten dimensions and compactifying down to seven dimensions. This would introduce a fourth mass parameter. On the IIB side, this would correspond to a compactification of the ten-dimensional theory on $T^3$ which simultaneously introduces a twist in the RR axion along with a twist in the torus geometry, as well as constant internal expectation values of the NS-NS and RR three-form field strengths. This is likely to be the most general massive seven-dimensional supergravity which can be produced using the axionic translation symmetries.

Many aspects of these results apply universally for generalized axionic reductions. Individually, the axionic symmetries will correspond to nilpotent generators of the global symmetry group. Hence the Scherk-Schwarz reduction will coincide with the simple linear ansatz (1) when an individual mass parameter is introduced. However, when several masses are simultaneously turned on, the reduction ansatz may involve quadratic and higher order terms as in Eq. (9). These terms result from the failure of the various nilpotent generators to commute with each other.

It would be interesting to investigate the interplay of U-duality with these Scherk-Schwarz reductions – some aspects of this issue have been addressed recently, in [12]. Introducing the mass parameters generically breaks some part of the global symmetry group which would otherwise appear in the compactified theory. However, it should be possible to write the massive theory in a U-duality invariant form, as long as the symmetry breaking parameters, i.e., the masses, are endowed with the appropriate transformation properties [13]. Thus, as is standard in spontaneous symmetry breaking, a broken symmetry will act as a transformation between distinct massive theories, or distinct “vacua” of the higher dimensional theory. In the present case, the full supergravity duality group in seven-dimensions is $SL(5,\mathbb{R})$. While we have argued that the mass parameters should form a representation of this group, we have only identified four such parameters for the seven-dimensional theory. Thus, the full massive theory must contain new masses beyond those considered here. In the context of the Scherk-Schwarz framework, it may be that the latter are associated with symmetries other than the axionic ones identified here, e.g., Eq. (7). Thus one probably has to extend the reduction ansatz to include more general global symmetries [1,6] to produce a U-duality invariant form. Another aspect of these constructions which would be interesting to study in the context of U-duality is the non-Abelian gauge symmetries which arise in the Scherk-Schwarz reductions [1,5].

Acknowledgements

N.K. was supported by NSF Grant PHY-9219345. R.R.K. was supported by a PPARC Advanced Fellowship and would like to thank the ITP and the Department of Physics at UCSB for their hospitality and where part of this work was completed under support by NSF Grant PHY94-07194. R.C.M. was supported by NSERC of Canada and by NSF Grant PHY94-07194.

References


BPS and non-BPS domain walls in supersymmetric QCD with $SU(3)$ gauge group

A.V. Smilga, A.I. Veselov

ITEP, B. Cheremushkinskaya 25, Moscow 117218, Russia

Received 27 February 1998
Editor: L. Alvarez-Gaume

Abstract

We study the spectrum of the domain walls interpolating between different chirally asymmetric vacua in supersymmetric QCD with the $SU(3)$ gauge group and including 2 pairs of chiral matter multiplets in fundamental and anti-fundamental representations. For small enough masses $m < m_+ = 0.286 A_{\text{QCD}}$, there are two different domain wall solutions which are BPS-saturated and two types of "wallsome sphalerons". At $m = m_+$, two BPS branches join together and, in the interval $m_+ < m < m_+ = 3.704 A_{\text{QCD}}$, BPS equations have no solutions but there are solutions to the equations of motion describing a non-BPS domain wall and a sphaleron. For $m > m_+$, there are no solutions whatsoever. © 1998 Published by Elsevier Science B.V. All rights reserved.

1. Introduction

Supersymmetric QCD is the theory involving a gauge vector supermultiplet $V$ and some number of chiral matter supermultiplets. The models of this class attracted attention of theorists since the beginning of the eighties and many interesting and non-trivial results concerning their non-perturbative dynamics have been obtained [1]. The dynamics depends in an essential way on the gauge group, the matter content, the masses of the matter fields and their Yukawa couplings.

The most simple in some sense variant of the model is based on the $SU(N)$ gauge group and involves $N - 1$ pairs of chiral matter supermultiplets $S_i \bar{S}_i$, $S_i^\alpha$, $\bar{S}_i^\alpha$ in the fundamental and anti-fundamental representations of the gauge group with a common mass $m$. The lagrangian of the model reads

$$\mathcal{L} = \left( \frac{1}{4 g^2} \text{Tr} \int d^2 \theta \ W^2 + \text{H.c.} \right)$$

$$+ \sum_{i=1}^{N-1} \left[ \frac{1}{4} \int d^2 \theta d^2 \bar{\theta} S_i \bar{S}_i + D_i S_i + D_i^\dagger \bar{S}_i \right]$$

$$- \frac{m}{2} \left( \int d^2 \theta S_i \bar{S}_i + \text{H.c.} \right).$$  

(1)

color and Lorentz indices are suppressed. In this case, the gauge symmetry is broken completely and the theory involves a discrete set of vacuum states. The presence of $N$ chirally asymmetric states has

0370-2693/98/$19.00 © 1998 Published by Elsevier Science B.V. All rights reserved.

Pii: S0370-2693(98)00401-8
been known for a long time. They are best seen in the weak coupling limit $m \ll \Lambda_{\text{SQCD}}$ where the chirally asymmetric states involve large vacuum expectation values of squark fields $\langle s_i \rangle \gg \Lambda_{\text{SQCD}}$ and the low energy dynamics of the model is described in terms of the colorless composite fields $s_i \Phi = 2 S_i S_j$. The effective lagrangian presents a Wess–Zumino model with the superpotential

$$\mathcal{W} = -\frac{2}{3} \frac{\Lambda_{\text{SQCD}}^{3N+1}}{m} - \frac{m}{2} \text{Tr} \Phi$$

(2)

The second term in Eq. (2) comes directly from the lagrangian (1) and the first term is generated dynamically by instantons. Assuming $s_i \Phi = X \delta_{ij}$ and solving the equation $\partial \mathcal{W} / \partial \chi = 0$ ($\chi$ is the scalar component of the superfield $X$), we find $N$ asymmetric vacua

$$\langle \chi \rangle_k = \left( \frac{4}{3} \frac{\Lambda_{\text{SQCD}}^{3N+1}}{m} \right)^{1/2} e^{\pi ik/N}$$

(3)

(the vacua $``k''$ and $``k + N''$ have the same value of the moduli $\langle \chi \rangle_k$ and are physically equivalent). These vacua are characterized by a finite gluino condensate

$$\langle \text{Tr} \lambda^2 \rangle_k = 8\pi^2 m \langle \chi^2 \rangle_k$$

(4)

It was noted recently [2] that on top of (3) also a chirally symmetric vacuum with the zero value of the condensate exists. It cannot be detected in the framework of Eq. (2) which was derived assuming that the scalar v.e.v. and the gluino condensate are nonzero and large, but is clearly seen if writing down the effective lagrangian due to Taylor, Veneziano, and Yankielowicz (TVY) [3] involving also the composite field

$$\Phi^3 = \frac{3}{32 \pi^2} \text{Tr} W^2$$

(5)

The corresponding superpotential reads

$$\mathcal{W} = \frac{2}{3} \Phi^3 \left[ \ln \frac{\Phi^3 \det \mathcal{M}}{\Lambda_{\text{SQCD}}^{3N+1}} - 1 \right] - \frac{m}{2} \text{Tr} \mathcal{M}$$

(6)

The presence of different degenerate physical vacua in the theory implies the existence of domain walls—static field configurations depending only on one spatial coordinate ($z$) which interpolate between one of the vacua at $z = -\infty$ and another one at $z = \infty$ and minimizing the energy functional. As was shown in [4], in many cases the energy density of these walls can be found exactly due to the fact that the walls present the BPS-saturated states.

The energy density of a BPS-saturated wall in SQCD with $SU(N)$ gauge group satisfies a relation [5]

$$e = \frac{N}{8\pi^2} \left[ \langle \text{Tr} \lambda^2 \rangle_+ - \langle \text{Tr} \lambda^2 \rangle_- \right]$$

(7)

where the subscript $\pm \infty$ marks the values of the gluino condensate at spatial infinities. Bearing Eqs. (7), (4) in mind, the energy densities of the BPS walls are

$$e_r = N \left( \frac{4m^{N-1}}{3} \right)^{1/N}$$

(8)

for the real walls and

$$e_c = 2e_r \sin \frac{\pi}{N}$$

(9)

for the complex walls. The RHS of Eqs. (7)–(9) presents an absolute lower bound for the energy of any field configuration interpolating between different vacua.

The relation (7) is valid assuming that the wall exists and is BPS-saturated. However, whether such a BPS-saturated domain wall exists or not is a non-trivial dynamic question which can be answered only in a specific study of a particular theory in interest. This question has already been studied in our previous works [5,7–9]. In particular, in [5,7,8] the simplest model of the class (1) with $N_c = 2, N_f = 1$ was analyzed. The results are the following:

1. For any value of the mass of the matter fields $m$, there are domain walls interpolating between a chirally asymmetric and the chirally symmetric vacua (we call them real walls). They are BPS-saturated.

2. A relation of this kind can be derived also for other variants of the theory involving exotic groups and more complicated matter content, but in general case the energy of a BPS wall and the gluino condensate are not related so directly [6].
2. There are also complex BPS solutions interpolating between different chirally asymmetric vacua. But they exist only if the mass is small enough \( m \leq m_* \approx 4.67059 \ldots A_{\text{SQCD}} \). When \( m > m_* \), BPS walls are absent.

3. In a narrow range of masses \( m_* < m \leq m_{**} \approx 4.83 A_{\text{SQCD}} \), complex domain walls still exist, but they are not BPS saturated anymore. At \( m > m_{**} \), there are no such walls whatsoever.

In Ref. [9], we studied the problem of existence of BPS-saturated domain walls in the model (1) with \( N \geq 3 \). The results are basically the same as for \( N = 2 \): the real walls exist for any \( m \) and are BPS-saturated, and there are two complex BPS branches which exist in a limited range \( m < m_0 \). The value of \( m_0 \) goes down with \( N \). For \( N = 3 \), \( m_0 = 0.28604 \ldots A_{\text{SQCD}} \) for \( SU(3) \), \( m_0 = 0.75589 \ldots A_{\text{SQCD}} \) for \( SU(4) \), etc. For large \( N \), \( m_0 \propto N^{-3} \).

The results concerning the BPS walls were obtained by solving numerically the first order BPS equations

\[
\partial_\chi \phi = e^{i\chi/2} \frac{\partial \phi}{\partial \phi}, \quad \partial_\chi \chi = e^{i\chi/2} \frac{\partial \chi}{\partial \chi} \tag{10}
\]

associated with the TVY lagrangian. The phase \( \delta \) depends on particular vacua between which the wall interpolates (see Refs. [5,10,9] for details).

To study the spectrum of the domain walls which are not BPS-saturated, one has to solve the equations of motion which are of the second order, and, technically, the problem is a little bit more involved. We did it earlier for \( N = 2 \) [8]. This paper is devoted to the numerical solution of the equations of motion for the \( SU(3) \) gauge group.

2. Solving equations of motion

The scalar potential corresponding to the superpotential (6) is

\[
U(\phi, \chi) = \left| \frac{\partial \phi}{\partial \phi} \right|^2 + \left| \frac{\partial \phi}{\partial \chi} \right|^2
= 4 \left| \phi^2 \ln \left( \phi^2 \chi^{2(N-1)} \right) \right|^2
+ (N-1)^2 \left| m \chi - \frac{4 \phi^3}{3 \chi} \right|^2 \tag{11}
\]

(from now on we set \( A_{\text{SQCD}} = 1 \)). The potential (11) has \( N + 1 \) degenerate minima. One of them is chirally symmetric: \( \phi = \chi = 0 \). There are also \( N \) chirally asymmetric vacua with \( \langle \chi \rangle_k \) given in Eq. (3) and

\[
\langle \phi \rangle_k = \left( \frac{3m}{4} \right)^{\frac{(N-1)/(3N)}{N}} \exp \left( -\frac{2i(N - 1)\pi k}{3N} \right) \tag{12}
\]

To study the domain wall configurations, we should add to the potential (11) the kinetic term which we choose in the simplest possible form

\[
\mathcal{L}_\text{kin} = |\partial \phi|^2 + |\partial \chi|^2 \tag{13}
\]

and solve the equations of motion with boundary conditions

\[
\phi(\pm \infty) = \langle \phi \rangle_0 = \rho_* \, , \quad \phi(\pm \infty) = R_* \exp \left( -2\pi i(N - 1) / 3N \right) \tag{14}
\]

\[
\chi(\pm \infty) = \langle \chi \rangle_0 = \rho_* \, , \quad \chi(\pm \infty) = \rho_* e^{i\tau/N} \tag{14}
\]

Thereby we are studying the walls interpolating between "adjacent" complex vacua. Actual calculations will be performed for \( N = 3 \) where all the vacua are adjacent. In principle, one could also study numerically the walls interpolating between the vacua \( k = 0 \) and \( k = 2 \), for say, \( SU(5) \) gauge group, etc. We expect the physical results for all such walls to be qualitatively the same.

It is convenient to introduce the polar variables \( \chi = \rho e^{i\alpha}, \phi = R e^{i\beta} \) after which the equations of motion acquire the form

\[
R'' - R \beta'^2 = 8R^3 \left[ L(L + 3/2) + \beta^2 \right] + (N - 1)^2 \left[ \frac{16R^3}{3 \rho^3} - 4mR^2 \cos(\beta) \right] \tag{15}
\]

\[
R \beta'' + 2R \beta' = 12R^2 \beta + 4(N - 1)^2 mR^2 \sin(\beta) \tag{15}
\]

\[
\rho'' - \rho \alpha'^2 = (N - 1) \frac{8R^4}{\rho} L + (N - 1)^2 \left[ m^2 p - \frac{16R^6}{9 \rho^3} \right] \tag{15}
\]
\[ \rho \alpha'' + 2 \rho' \alpha' = (N - 1) \frac{8 R^4}{\rho} \beta, \]
\[ - (N - 1)^2 \frac{8 m R^3}{3 \rho} \sin(\beta_+), \quad (15) \]

where \( L = \ln[R^2 \rho^{2(N-1)}]. \) \( \beta_+ = 3 \beta + 2(N - 1) \alpha, \)
\( \beta_- = 3 \beta - 2 \alpha, \) with the boundary conditions
\[ \rho(-\infty) = \rho(\infty) = \rho_+; \quad R(-\infty) = R(\infty) = R_+; \]
\[ \alpha(-\infty) = \beta(-\infty) = 0; \quad \alpha(\infty) = \pi/N; \]
\[ \beta(\infty) = - \frac{2(N-1)\pi}{3N} \]
\[ (16) \]

When \( N = 2, \) the system (15) is reduced to that studied in Ref. [8]. The system (15) involves one integral of motion
\[ T - U = R^2 + \rho'^2 + R^2 \beta'^2 + \rho^2 \alpha'^2 - 4 R^4 (L^2 + \beta^2) \]
\[ - (N - 1)^2 \left[ \frac{16 R^6}{9 \rho^2} - \frac{8 m R^3}{3} \cos(\beta_+) \right] \]
\[ = \text{const} \quad (17) \]

In our case, \( \text{const} = 0 \) due to boundary conditions (16). The phase space of the system (15) is 8-dimensional and a general Cauchy problem involves 8 initial conditions. The problem is simplified, however, when noting that the wall solution should be symmetric with respect to its center. Let us seek for the solution centered at \( z = 0 \) so that
\[ \rho(z) = \rho(-z), \quad R(z) = R(-z), \]
\[ \alpha(z) = \pi/N - \alpha(-z) \]
\[ \beta(z) = -2(N - 1) \pi/(3N) - \beta(-z) \quad (18) \]

Indeed, one can be easily convinced that the Ansatz (18) goes through the Eqs. (15). It is convenient to solve the Eqs. (15) numerically on the half-interval from \( z = 0 \) to \( z = \infty. \) The symmetry (18) dictates \( \rho'(0) = R'(0) = 0, \alpha(0) = \pi/(2N), \beta(0) = - (N - 1) \pi/(3N) \) which fixes 4 initial conditions. Four others satisfy the relation (17). Thus, we are left with 3 free parameters, say, \( \rho(0), R(0), \) and \( \beta'(0), \) which should be fitted so that the solution approach the complex vacuum in Eq. (14) at \( z \rightarrow \infty. \)

All the solutions obtained with such a procedure are presented in Fig. 1 for the parameter \( R(0), \)

![Fig. 1. The ratio \( \eta = R(0)/R_0(0) \) for the solutions of the equations of motion as a function of mass for the \( SU(3) \) theory. The solid lines describe the BPS solutions and the dashed lines describe the non-BPS wall and the sphalerons.](image-url)
barrier separating the lower BPS branch and the configuration of two distant real walls at \( R(0) = 0 \).

At \( m = m_* \), two BPS minima fuse together and the energy barrier separating them disappears. The upper sphaleron branch coincides with the BPS solution at this point. When \( m \) is increased above \( m_* \), the former BPS minimum is still a minimum of the energy functional, but its energy is now above the BPS bound (see Fig. 2b). The corresponding solution is described by the analytic continuation of the upper sphaleron branch. The lower dashed branch in the region \( m_* < m < m_{**} \) is still a sphaleron. At the second critical point \( m = m_{**} \), the picture is changed again (see Fig. 2c). The local maximum and the local minimum fuse together and the only one remaining stationary point does not correspond to an extremum of the energy functional anymore. At larger masses, no non-trivial stationary points are left.

Our findings are illustrated in Figs. 3 and 4 where the energies of the non-BPS wall and of the lower sphaleron branch are plotted as a function of \( m \). In Fig. 3, the ratios of the energies of both branches to the BPS bound (9) are plotted. The lower line in Fig. 3 corresponds to the stable wall solution and the upper line to the sphaleron branch. For almost all \( m < m_* \), the wall solution is globally stable. When \( m_* - m \) becomes very small, it is stable only locally: we see that, at \( m = m_* \), where two branches are fused together, their energy exceeds slightly the energy of two real walls \( 2 \epsilon_r \).

In Fig. 4 the sphaleron energy is redrawn in logarithmic scale in the units of \( 2 \epsilon_r \). Unfortunately, we do not have good numerical data at \( m \leq 0.7 \)—our relative uncertainty becomes large. We see, however, that the logarithmic plot in Fig. 4 is pretty much linear, the best fit is

\[
\kappa'(m) = 5.49 \cdot 10^{-3} \exp(2.11m) \tag{19}
\]

The fit (19) cannot be valid for very small masses: we expect \( \kappa'(0) = 0 \) which means that the straight

---

In contrast to the case \( N = 2 \), the existence of such a barrier could not be established from the BPS spectrum alone. The matter is that, while \( 2 \epsilon_r = \epsilon_s \) for \( N = 2 \) and the presence of the maximum is guaranteed by the Roll theorem, \( 2 \epsilon_r > \epsilon_s \) in our case and one could in principle imagine a situation where \( E \) falls down monotonically when \( R(0) \) is increased from zero up to its value at the lower BPS branch. As will be discussed later, our numerical results are not good enough to make a rigid statement on the form of the function \( E[R(0)] \) at small \( R(0) \) in the region \( m < m_* \), but they strongly suggest that the energy barrier (though a tiny one) is present.
Fig. 5. Profile $R(z)$ for the lower sphaleron branch at $m = .7$.

line in Fig. 4 should bend down at small enough masses due to a preexponential factor $\sim m^\alpha$ which we cannot determine from our data. Anyway, it is seen from Fig. 4 that $\ln k^X$ does not "want" to hit infinity at a finite mass, and we assume that it does not (though we cannot exclude it as a logical possibility). In terms of Fig. 2, that means that the energy barrier on the left (for illustrative purposes, it is very much exaggerated) is present for all masses.

Finally, we present for illustration the profile $R(z)$ for the lower sphaleron branch at $m = .7$ (Fig. 5). As was expected, it resembles very much a combination of two separate real domain walls.

3. Discussion

Our main conclusion is that, besides the critical mass $m_s$ beyond which BPS solutions disappear, also a second critical mass $m_{c,s}$ exists beyond which no complex wall solution can be found whatsoever. This was the case for $SU(2)$ [8] and, as we see now, this is also the case for $SU(3)$. Seemingly, the same situation holds for any $N$. That means in particular that no domain walls connecting different chirally asymmetric vacua are left in the pure supersymmetric Yang–Mills theory corresponding to the limit $m \to \infty$, and only the real domain walls connecting the chirally symmetric and a chirally asymmetric vacuum states survive in this limit. That contradicts an assumption of Ref. [11] that it is complex rather than real domain walls which are present in the pure SYM theory (Witten discussed them in the context of brane dynamics).

One has to make a reservation here: our result was obtained in the framework of the TVY effective lagrangian (6) whose status [in contrast to that of the lagrangian (2)] is not absolutely clear: the field $\Phi$ describes heavy degrees of freedom (viz. a scalar glueball and its superpartner) which are not nicely separated from all the rest. However, the form of the superpotential (6) and hence the form of the lagrangian for static field configurations is rigidly dictated by symmetry considerations; the uncertainty involves only kinetic terms. It is reasonable to assume that, as far as the vacuum structure of the theory is concerned (but not e.g. the excitation spectrum—see Ref. [8] for detailed discussion), the effective TVY potential (6) can be trusted. A recent argument against using Eq. (6) that the chirally symmetric phase whose existence follows from the TVY lagrangian does not fulfill certain discrete anomaly matching conditions [12] is probably not sufficient. First, it assumes that the excitation spectrum in the symmetric phase is the same as it appears in the TVY lagrangian which is not justified. Second, it was argued recently that the TVY lagrangian describes actually all the relevant symmetries of the underlying theory and the absence of the anomaly matching is in a sense an "optical illusion" [13].

The main distinction of the $SU(3)$ case considered here compared to the $SU(2)$ theory is that the values of two critical values are rather different (in $SU(2)$ case they were pretty close: $m_s = 4.67059\ldots$ and $m_{c,s} = 4.83$). This is due to the fact that the energy of the complex BPS wall $e_s$ is less in this case than the energy of two real walls $2e_r$. When we increase the mass and go above $m_s$, the energy of the wall first has to rise from $e_s$ to $2e_r$. Only then the complex domain wall "bound state" can break apart into its "constituents", the real walls. \footnote{Actually, as we have seen, the energy barrier in Fig. 2 does not allow the complex wall to break apart until its energy goes a little bit above the limit $2e_r$.}

One can expect that $m_s$ and $m_{c,s}$ differ more and more as $N$ grows. A tentative guess is that $m_{c,s}$.

\footnote{Actually, as we have seen, the energy barrier in Fig. 2 does not allow the complex wall to break apart until its energy goes a little bit above the limit $2e_r$.}
is roughly $N$-independent (to be compared with $m_1(N) \propto N^{-3}$). Of course, that can be confirmed or disproved by only actual numerical study. Note, however, that numerical calculations become more and more difficult as $N$ grows—the instabilities characterized by eigenvalues of the Jacobi matrix of the system (15) near the minima (3), (12) grow as $N^2$.

Acknowledgements

This work was supported in part by the RFBR-IN-TAS grants 93-0283, 94-2851, 95-0681, and 96-370, by the RFFI grants 96-02-17230, 97-02-17491, and 97-02-16131, by the RFBR-DRF grant 96-02-00088, by the U.S. Civilian Research and Development Foundation under award # RP2-132, and by the Schweizerischer National Fonds grant # 7SUPJ048716.

References

Limits on $\tan \beta$ in SU(5) GUTs with gauge-mediated supersymmetry breaking

H. Hamidian $^a$, K. Huitu $^b$, K. Puolamäki $^b$, D.-X. Zhang $^b$

$^a$ Department of Physics, Stockholm University, Box 6730, S-113 85 Stockholm, Sweden
$^b$ Helsinki Institute of Physics, P.O. Box 9, FIN-00014 University of Helsinki, Finland

Received 3 March 1998; revised 31 March 1998
Editor: P.V. Landshoff

Abstract

By considering the constraints from nucleon decay we obtain upper limits on $\tan \beta$ in generalized supersymmetric SU(5) grand unified theories with gauge-mediated supersymmetry breaking. We find that the predicted values of $\tan \beta$ in these models are mostly inconsistent with the constraints from nucleon decay.

PACS: 11.30.Pb; 12.10.Dm; 12.60.Jv; 14.80.Ly

Supersymmetric (SUSY) grand unified theories (GUTs) are presently considered to be among the most promising candidates for physics beyond the standard model (SM). However, for phenomenological reasons, supersymmetry cannot be exact and it is usually assumed that the theory includes a visible sector containing the observable particles, and a hidden sector where supersymmetry is broken. SUSY breaking can be communicated to the visible sector either by gravitational interactions, as in supergravity (SUGRA) inspired models, or by SM gauge interactions, as in theories with gauge-mediated SUSY breaking (GMSB).

GMSB models were initially studied in the early 1980’s [1] and have recently become the subject of much theoretical investigation. The revival of interest in these models [2] is largely due to recent dramatic improvements in our understanding of non-perturbative effects in SUSY gauge theories as a result of the pioneering works of Seiberg [3] and Seiberg and Witten [4]. Many new mechanisms for dynamical SUSY breaking (DSB) have been found since the appearance of [3,4] and new DSB models have been constructed [5]. From a phenomenological point of view, GMSB theories are interesting for a number of reasons. In these theories gauge interactions provide flavor-symmetric SUSY breaking terms and thus naturally suppress the flavor-changing neutral currents associated with soft squark and slepton masses. They also predict approximately degenerate squark and slepton masses (see below) and are specified by a relatively small number of parameters.

In GMSB theories the $SU(3) \times SU(2) \times U(1)$ gauge interactions of the “messenger” fields communicate SUSY breaking from a hidden sector to the fields of the visible world. 1

1 Here we do not consider the modification of direct mediation models such as the one proposed, e.g., in [6].
models [2], in addition to the particles in the minimal supersymmetric standard model (MSSM), there exists at least one singlet superfield \( S \) which couples to vector-like messenger superfields \( V + \nabla \) through the superpotential interaction
\[
W_{\text{mresse}} = \lambda_v SV\nabla. \tag{1}
\]

At a scale \( \Lambda \sim 10^{10} \text{ TeV} \), which is not much higher than the weak scale, SUSY is broken and both the lowest and \( F \)-component, \( F_5 \), of the singlet superfield \( S \) acquire vacuum expectation values (VEVs) through their interactions with the hidden sector. The VEV, \( \langle S \rangle \), gives masses to the vector-like supermultiplets \( V + \nabla \), while \( \langle F_5 \rangle \) induces mass splittings within the supermultiplets. As a result, the gauginos and sfermion masses are generated through their gauge couplings to the messenger fields. The gauginos receive masses at one-loop, \( m_a \sim (\alpha/4\pi) \Lambda \), where \( \Lambda = \langle F_5 \rangle/\langle S \rangle \), while squarks and sleptons do so only at two-loop order, \( \tilde{m}^2 \sim (\alpha/4\pi)^2 \Lambda^2 \). This implies that \( m_a \sim \tilde{m} \), which is one of the attractive features of GMSB theories.

In the minimal version of the GMSB model [2], the messenger fields belong to the \( 5 + 5 \) or \( 10 + 10 \) representations of the \( SU(5) \) gauge group, and the messenger Yukawa couplings, \( \lambda_V \)’s in (1), in any given \( SU(5) \) representation are taken to be equal at the unification scale \( M_{\text{GUT}} \). Consequently, the spectrum at the messenger scale consists of a set of fields in complete \( SU(5) \) representations and the mass splitting among the fields in a representation is induced through the renormalization group running of the messenger Yukawa couplings from \( M_{\text{GUT}} \) down to the messenger scale \( \Lambda_m \).

While the phenomenological implications of the minimal GMSB model have been extensively studied [7–9], non-minimal generalizations of this class of theories have seen less investigation. In this Letter, we study generalized models of GMSB proposed by Martin [10] in which the messenger fields do not necessarily form complete \( SU(5) \) GUT multiplets. It is, in fact, not difficult to see how one may be naturally led to consider messenger fields which belong to incomplete representations of the \( SU(5) \) gauge group in the generalized GMSB models. This is because the unification of the messenger Yukawa couplings at the GUT scale – whose MSSM analogue is the so-called \( b - \tau \) unification [11] – is not necessarily required for gauge unification. Suppose, for example, that in addition to \( S \) there exist singlet superfields, \( S' \), whose VEVs (but not the VEVs of their \( F \)-components) are just below the GUT scale, and which couple only to some components of an \( SU(5) \) multiplet 2. Then within the \( SU(5) \) multiplet these superfields acquire masses of order \( \sigma (M_{\text{GUT}}) \) and decouple from the low-energy spectrum. The other components, which get their masses only through couplings with the superfield \( S \), obtain masses of order \( \lambda \langle S \rangle \sim A_m \). Since \( \langle F_5 \rangle \) is much smaller than the masses of the heavy superfields, these (missing) particles make negligible mass contributions and play a less important role in determining the MSSM mass spectrum.

A fruitful approach for examining the phenomenological viability of SUSY GUTs has been to study processes that contribute to the nucleon decays [13]. However, in contrast to the minimal GMSB theories, in which constraints from the nucleon decay yield quite strong results [9], the situation in this regard is somewhat more complicated in the generalized GMSB models. In the minimal model one begins with the renormalization group running of the SM gauge coupling constants from the electroweak scale up to the GUT scale to determine the mismatch between the SM and the \( SU(5) \) gauge coupling constants at the GUT scale. The mismatch, which is expressed in terms of the sum of the contributions coming from threshold corrections at the weak scale, the messenger scale, and the GUT scale, can then be used to calculate the masses of the color-triplet Higgs bosons, \( M_{\tilde{H}_c} \). On the other hand, in the generalized GMSB models the mass splitting within a given \( SU(5) \) representation is undetermined and may result in large contributions to the mismatch at the GUT scale. This means that the masses of the color-triplet Higgs bosons cannot, in general, be reliably determined in the same way as in the minimal model.

2Horizontal symmetries, e.g., can be used to construct such models [12].
However, it is still possible to obtain useful constraints in the generalized SU(5) GMSB models. The color-triplet Higgs superfields belong to the SU(5) representation of SU(5) fields in the messenger sector of the generalized GMSB model. Let us now describe the GMSB models that are the subject of this Letter. Following Martin [10], we consider five possible types of chiral superfields in the messenger sector.

\[ n_L: L + \overline{L} = (1,2,-\frac{1}{2}) + \text{conj.}, \]

\[ n_D: D + \overline{D} = (3,1,\frac{1}{2}) + \text{conj.}, \]

\[ n_E: E + \overline{E} = (1,1,1) + \text{conj.}, \]

\[ n_U: U + \overline{U} = (3,1,-\frac{2}{3}) + \text{conj.}, \]

\[ n_Q: Q + \overline{Q} = (3,2,\frac{1}{6}) + \text{conj.}, \]

where the multiplicities of the messenger fields are denoted by \( n_L, n_D, n_E, n_U, n_Q \).

Requiring that the gauge couplings remain perturbative, and assuming messenger field masses that do not greatly exceed \( 10^4 \) TeV, leads (see [10] for further discussion) to the following set of multiplicities for the messenger fields

\[ (n_L, n_D, n_E, n_U, n_Q) \leq \begin{array}{c} (1,2,2,0,1) \\ (1,1,1,1,1) \\ (1,0,0,2,1) \\ (4,4,0,0,0). \end{array} \]

The general low energy superpotential of the messenger sector is

\[ W_{\text{mess}} = \sum_{n_L} \lambda^I_L \overline{SU} \overline{\mathcal{L}} + \sum_{n_D} \lambda^I_D S \overline{D} \overline{T} + \sum_{n_E} \lambda^I_E S \overline{E} \overline{\mathcal{E}} 
+ \sum_{n_U} \lambda^I_U \overline{S} \overline{U} \overline{T} + \sum_{n_Q} \lambda^I_Q S \overline{Q} \overline{\mathcal{E}}. \]

and the MSSM spectrum can be determined once the messenger sector is fixed. In the numerical estimates that are reported here, we shall use \( A_0 = 10^4 \) TeV.

We can now directly calculate the nucleon decay rates for the GMSB models listed above by using the process \( n \rightarrow K^0 \overline{\nu}_\mu \) as the characteristic mode. The short- and long-distance corrections and the hadronic matrix elements are taken at their conservative values, as e.g. in [9], giving a lower bound on the parameter \( M_{H_c} \sin 2 \beta \). We have studied all the 53 possible models which satisfy criteria (3) and contain massive gauginos. By setting the upper limit on the mass of the triplet Higgs at \( 10^{17} \) GeV, upper bounds on \( \tan \beta \) can be obtained and some interesting configurations for \( \Lambda = \langle F_T \rangle / \langle S \rangle = 100 \) TeV are listed in Table 1. The parameters in the columns are the ones needed in the nucleon decay formula. We find a general bound, \( \tan \beta < 10 \), except in cases (1,4,0,0,0), (2,4,0,0,0), and (1,3,0,0,0) for which this bound is \( \tan \beta < 17 \).

There is also a lower bound on \( \tan \beta \) following from the nucleon decay constraints. However, this constraint is less severe than the one obtained by requiring that the Yukawa couplings should not blow up at the GUT scale, giving \( \tan \beta > 0.85 \) [9].

For comparison, we have also calculated the value of \( \tan \beta \) with the assumption of radiatively-broken SU(2) \( \times U(1) \) symmetry when trilinear and bilinear soft couplings vanish at the messenger scale - which, among other things, free these models from the supersymmetric CP problem and make them extremely predictive. For the computations we use the full one-loop effective potential [8]. The calculated values of \( \tan \beta \) are plotted versus the upper limit from nucleon decay in Fig. 1. The calculations are done for \( \Lambda = 100 \) TeV (black circles) and \( \Lambda = 200 \) TeV (gray circles).

Table 1

Limits on \( \tan \beta \) with \( \Lambda = 100 \) TeV

\[
\begin{array}{cccccc}
(n_L, n_D, n_E, n_U, n_Q)^* & (\tan \beta)_{\text{min}}^b & (\tan \beta)_{\text{max}}^a & (\tan \beta)_{\text{LSP}} & (\tan \beta)_{\text{Yuk}} & (\tan \beta)_{\text{GUT}} \\
(1,4,0,0,0)^c & 17 & 301 & 263 & 2565 & 2262 & 373 \\
(1,3,0,0,0)^c & 13 & 245 & 263 & 1971 & 1894 & 369 \\
(2,4,0,0,0)^c & 11 & 384 & 521 & 2565 & 2287 & 527 \\
(1,1,0,0,0)^d & 5 & 301 & 263 & 724 & 1006 & 383 \\
\end{array}
\]

* \( n_i \) specify the numbers of different types of messenger fields as defined in Eq. (2). b Limit assumes \( M_{H_c} \leq 10^{17} \) GeV. c These three configurations are the only choices which allow \( \tan \beta > 10 \). d 5+3 model.
Fig. 1. The upper limit on $\tan \beta$ from nucleon decay vs. $\tan \beta$ calculated by assuming radiatively-broken $SU(2) \times U(1)$ symmetry and vanishing bilinear and trilinear soft couplings at the messenger scale. The black circles correspond to $\Lambda = 100$ TeV and the open rectangles to $\Lambda = 200$ TeV.

As serious SUSY GUT candidates beyond the standard model, the bilinear and/or trilinear soft terms cannot vanish at the messenger scale, and gauge groups other than $SU(5)$ – together with their associated implementation of dynamical SUSY breaking – are required for acceptable low-energy phenomenology.

Acknowledgements

H.H. thanks the Swedish Natural Science Research Council for financial support. The work of K.P. is supported by a grant from the Magnus Ehrnrooth Foundation, and the work of K.H., K.P. and D.X.Z. by the Academy of Finland (No. 37599).
References


Spaces with torsion from embedding, and the special role of autoparallel trajectories

Hagen Kleinert 1, 2, Sergei V. Shabanov 3, 4
Institute for Theoretical Physics, FU-Berlin, Arnimallee 14, D-14195 Berlin, Germany

Received 21 February 1998
Editor: P.V. Landshoff

Abstract

As a contribution to the ongoing discussion of trajectories of spinless particles in spaces with torsion we show that the geometry of such spaces can be induced by embedding their curves in a euclidean space without torsion. Technically speaking, we define the tangent (velocity) space of the embedded space imposing non-holonomic constraints upon the tangent space of the embedding space. Parallel transport in the embedded space is determined as an induced parallel transport on the surface of constraints. Gauss’ principle of least constraint is used to show that autoparallels realize a constrained motion that has a minimal deviation from the free, unconstrained motion, this being a mathematical expression of the principle of inertia. In contrast, geodesics play no special role in the constrained dynamics, making them less likely candidates for particle trajectories.

1 On an affine manifold equipped with a metric, there exist two preferred connections compatible with the metric [1]. One is the Riemann connection defined only by the metric. In a coordinate basis in the tangent space of the manifold, the coefficients of this connection are Christoffel symbols

\[ \Gamma_{\mu\nu\kappa} = \frac{1}{2} \left( g_{\mu\lambda,\nu} + g_{\nu\lambda,\mu} - g_{\mu\nu,\lambda} \right). \]

Here \( g_{\mu\nu} \) are components of the metric tensor, and indices after a comma stand for the corresponding derivatives, \( T_{\mu\nu\ldots\kappa} \ldots = \partial_{\lambda} \partial_{\mu} \cdots T_{\mu\nu\ldots} \cdots \). By definition, the covariant derivative formed with this Riemann connection satisfies the metricity condition

\[ \bar{D}_\mu g_{\nu\kappa} = \partial_\mu g_{\nu\kappa} - \Gamma^\lambda_{\mu\nu} g_{\lambda\kappa} - \Gamma^\lambda_{\mu\kappa} g_{\nu\lambda} = 0. \] (2)

Apart from \( \Gamma^\lambda_{\mu\nu} \), there exists also a Cartan connection \( \Gamma^\lambda_{\mu\nu} \). It satisfies the same compatibility condition with the metric:

\[ D_\mu g_{\nu\kappa} = \partial_\mu g_{\nu\kappa} - \Gamma^\lambda_{\nu\kappa} g_{\lambda\mu} - \Gamma^\lambda_{\kappa\mu} g_{\nu\lambda} = 0. \] (3)

It can always be represented in the form [1]

\[ \Gamma_{\mu\nu\kappa} = \frac{\mu}{\nu\kappa} + K_{\mu\nu\kappa}, \] (4)

where \( K_{\mu\nu\kappa} \) is any antisymmetric tensor in \( \nu\kappa \), called the contorsion tensor [1]

\[ K_{\mu\nu\kappa} = S_{\mu\nu\kappa} - S_{\nu\mu\kappa} + S_{\kappa\mu\nu}, \] (5)
where \( S_{\mu \nu} = g_{\kappa \lambda} S_{\mu}^\kappa S_{\nu}^\lambda \) and \( S_{\mu \nu}^\lambda \) is the torsion tensor.

\[
S_{\mu \nu}^\lambda = \frac{1}{2} \{ \Gamma_{\mu \nu}^\lambda - \Gamma_{\nu \mu}^\lambda \}.
\]  

(6)

Each connection compatible with the metric on the manifold defines a curve which parallel-transports its tangent vector along itself. Let \( v^\mu \) stand for the tangent vector and \( \dot{v}^\mu \) for its derivative with respect to the affine parameter, the proper time on the curve. Then the equation for the curve which parallel-transports its tangent vector along itself with respect to the Cartan connection reads

\[
\frac{Dv^\mu}{dt} \equiv \dot{v}^\mu + \Gamma_{\lambda \nu}^\mu v^\lambda \epsilon^\nu = 0.
\]  

(7)

It describes autoparallels, the straightest curves in Riemann-Cartan space. The reason for this name will be explained later in Section 5. The same equation with the Christoffel symbols describes geodesics, the shortest curves with respect to the metric \( g_{\mu \nu} \).

For the Cartan connection (4), the deviation from the geodesics is caused by a torsion force in (7) coming from the symmetrical part of the contorsion tensor \( K_{(\mu \nu)} v^\mu \epsilon^\nu = -2S_{\mu \nu}^\lambda v^\mu \epsilon^\nu \).

Apart from extremizing a length between two fixed endpoints, geodesics in a Riemannian space can be obtained by embedding the Riemannian space in a euclidean space of a higher dimension. This is done by imposing certain constraints on the Cartesian coordinates spanning the euclidean space. The points on the constraint surface constitute the embedded Riemannian space. Straight lines in the euclidean space, which are geodesic and autoparallel and also determine a free motion in that space, become geodesics when the motion is restricted to the constraint surface. The restriction of the free motion to the constraint surface is done in a conventional way, i.e., by adding the equations of constraints to the equations of motion. When the constraint force is removed, geodesic trajectories turns into straight lines in the embedding space.

For dynamics in Riemann-Cartan spaces, no such embedding is known. The purpose of this letter is to fill this gap. The new embedding will have the property that autoparallels are realized as trajectories of a constrained free motion.

Key observation of our theory is the fact that in order to define a geometry (metric and the law of parallel-transport) by embedding it is not necessary to constrain the positions in a euclidean space. One may impose constraints only on the tangent (or velocity) space, thereby defining a physical tangent space as a subspace embedded in a bigger velocity space. This is sufficient to define all curves in the embedded space as a special subset of all curves in the embedding space because each curve is specified by its tangent vector. The metric in the embedded space is naturally induced by restricting the scalar product in the bigger (embedding) tangent space to the constraint surface. The induced connection is uniquely determined by the compatibility condition of the embedding of the tangent space with the parallel-transport law in the embedding space. This means the following: Take a curve in the original space connecting points 1 and 2. This curve is then embedded into a bigger euclidean space by specifying the tangent vector of the image curve. A vector from the tangent space at point 1 is parallel-transported along the curve to point 2, and then it is embedded in the bigger space. We require that the resulting vector must be the same as the one obtained in the opposite way: The vector at point 1 is first embedded into a bigger euclidean space and then parallel-transported along the image of the curve connecting points 1 and 2. This compatibility condition ensures that the connection in the original space is uniquely determined by the embedding law.

Constraints imposed on the tangent space can be non-holonomic, and this is the source of torsion. The notion of "holonomic" and "non-holonomic" constraints is the same as in classical mechanics. For a mechanical system, generalized velocities are elements of the tangent space of its configuration space. Let the motion be subject to constraints linear in velocities. According to the Hertz classification [2], constraints are said to be holonomic if they are integrable (i.e., equivalent to some constraints on the configuration space only), and non-holonomic if they are non-integrable. Sometimes dynamical systems with non-holonomic constraints are simply called non-holonomic systems. It is important to realize that the motion of non-holonomic systems does not occur on any submanifold of the configuration space, nonetheless it is described by a less number of parameters than the corresponding unconstrained motion.
Upon an embedding via non-holonomic constraints on the tangent space, any curve that parallel-transports its tangent vector along itself has an image with same property. Therefore straight lines, which are autoparallels and geodesics with respect to a trivial euclidean connection, are natural images of autoparallels in the embedded space. Using Gauss’ principle of least constraint we show that autoparallels describe a constrained motion such that the acceleration (or the force) induced by the constraints has a least deviation from the acceleration of the corresponding unconstrained motion, while geodesics play no special role in the constrained dynamics. This comprises the main result of our paper.

Note that the theory of general relativity considered here admits torsion in the geometry of spacetime, as in the Einstein-Cartan theory [4]. If general coordinate invariance is the only principle for finding equations of motion, both geodesics and autoparallels are equally good candidates for trajectories of spinless test particles [5,6]. Recently, however, convincing physical arguments have been put forward, based on the analogy of torsion with dislocations in crystals and their effect on the geometry of parallelograms (closure failure) [7], which tend to favor autoparallel trajectories.

An important guide for finding the correct laws of interaction between matter and spacetime are gauge principles and minimal coupling [8], according to which all derivatives in a Lagrangian of a free theory should be replaced by the covariant derivatives in order to obtain the interacting theory. In such an approach, scalar fields are decoupled from torsion, in which case geodesics would be the true trajectories of spinless point particles. This is a consequence of the minimality of the coupling. There is no problem with the gauge principles. Eq. (7) is just as gauge covariant as the geodesic equation \( D_{\alpha}^{\mu} / dt = 0 \).

In this note we want to find more hints on the correct dynamics of masspoints with the help of another geometric consideration. For this we remember that the mathematical technique of embedding has been a powerful tool in studying geometrical forces in non-euclidean configuration spaces. Our approach based on the constrained dynamics might be useful in constructing possible field models of interaction between matter and a generic spacetime connection where autoparallels are true trajectories of spinless particles. Up to now neither experiments nor theoretical principles forbid such theories.

2. Let \([x]\) be a Euclidean space and \(x', i = 1, 2, \ldots, M = \text{dim}[x]\), be a set of coordinates. Let \([q]\) be a space of a smaller dimension, \(N = \text{dim}[q]\), spanned by local coordinates \(q^\mu, \mu = 1, 2, \ldots, N\). Tangent spaces of \([x]\) and \([q]\) are denoted as \(T[x]\) and \(T[q]\), respectively. Consider a curve \(q^\mu(t)\) in \([q]\) and tangent spaces at each point of the curve, \(T[q(t)]\). We define a curve \(x'(t)\) in the embedding space \([x]\) by specifying its tangent vector

\[ v' = e'_\mu(q) v^\mu, \quad v^\mu = \dot{q}^\mu(t), \quad v' = \ddot{x}'(t), \]  

where coefficients \(e'_\mu(q)\) are some functions on \([q]\). From (8) follows

\[ x'(t) = x'(0) + \int_0^t dq^\mu e'_\mu(q). \]

For any curve \(q^\mu(t)\) in \([q]\), Eq. (9) determines a curve in the space \([x]\) up to a global translation on a vector \(x'(0)\). Thus it determines an embedding of the space of all paths in \([q]\) into the space of all paths in \([x]\). Assuming Eq. (8) to hold for all curves in the space \([q]\) passing through some point \(q^\mu\), we specify the embedding of \(T[q]\) at this point into the Euclidean space \(T[x]\). Note that tangent spaces at all points of \([x]\) are the same and coincide with the space \([x]\) because the parallel transport is trivial in the space \([x]\) (the connection \(G^i_{jk}\) vanishes identically). So, the shift of the image curve \(x'(t)\) on a constant vector in (8) is irrelevant for the embedding of the tangent space. The embedding of the path space or the tangent space includes the case of the pointwise embedding of the space \([q]\) itself into \([x]\), but it appears to be more general as we shall see.

Consider two sets of tangent spaces \(T[q(t)]\) and \(T[x(t)]\) at points of the curve \(q^\mu(t)\) and of its image \(x'(t)\) defined by (9). We embed the space \(T[q(t)]\) in \(T[x(t)]\) so that for any element \(T^\nu\) of \(T[q(t)]\), its image in \(T[x(t)]\) is

\[ T' = e'_\nu(q(t)) T^\nu. \]

We perform a parallel transport of the image vector \(T'\) along the curve \(x'(t)\). Since the connection in the space \([x]\) is zero, an infinitesimal change of \(T'\) is
determined by the total derivative with respect to the affine parameter

\[
\frac{dT^i}{dt} = v^i D T^i = \frac{dT^i}{dt}. \tag{11}
\]

Similarly, an infinitesimal change of the vector \( T^\mu \) under the parallel transport along the curve \( q^\mu(t) \) is specified by the covariant derivative

\[
\frac{DT^\mu}{dt} = v^\nu D T^\nu = v^\nu \left( \partial_v T^\mu + \Gamma^\mu_{\nu\lambda} T^\lambda \right), \tag{12}
\]

with \( \Gamma^\mu_{\nu\lambda} \) being the connection on the space \([q]\). Now we come to the crucial condition for our embedding procedure: We require that the vector \( DT^\mu/dt \) must be the image of the vector \( DT^\mu/dt \), that is,

\[
\frac{DT^\mu}{dt} = \epsilon^\mu_{\nu} \frac{DT^\nu}{dt}. \tag{13}
\]

Eq. (13) has a transparent geometrical meaning. The parallel transport of the vector \( T^\mu \) along any curve \( q^\mu(t) \) and the subsequent embedding of the resulting vector into the bigger space \( T[x] \) via the relation (10) gives an element of \( T[x] \). This element must coincide with the one obtained in the opposite way in which the vector \( T^\mu \) is first embedded and then parallel-transported along the image \( x^i(t) \) of the curve \( q^\mu(t) \). This implies that the embedding of the tangent space \( T[q] \) at any point of \([q]\) is compatible with the parallel transport on \([q]\). The compatibility condition (13) uniquely determines the connection coefficients \( \Gamma^\mu_{\nu\lambda} \) via the embedding coefficients \( \epsilon^\mu_{\nu} \).

Before we proceed to prove this statement, let us introduce some useful notations. For any two vectors from \( T[x] \), one can introduce a scalar product associated with the Cartesian metric on \([x]\)

\[
(\tilde{T}, T) = \delta^\nu_{\mu} \tilde{T}^\nu T^\mu. \tag{14}
\]

If the vectors \( T^i \) and \( \tilde{T}^i \) are the images of \( T^\mu \) and \( \tilde{T}^\mu \), respectively, then the embedding coefficients \( \epsilon^\mu_{\nu} \) determine an induced metric on \([q]\)

\[
(\tilde{T}, T) = g_{\mu\nu} \tilde{T}^\nu T^\mu, \quad g_{\mu\nu} = (\epsilon_{\mu}, \epsilon_{\nu}). \tag{15}
\]

It is useful to introduce the quantity

\[
\epsilon^{i\mu} = \epsilon^i_{\nu} g^{\nu\mu}, \tag{16}
\]

where \( g^{\mu\lambda} g_{\lambda\nu} = \delta^{\mu}_{\nu} \). From (15) follows that

\[
(\epsilon^{\mu}, \epsilon^{\nu}) = g^{\mu\nu}, \quad (\epsilon^\mu_{\nu}, \epsilon_{\nu}) = \delta^\mu_{\nu}. \tag{17}
\]

Assuming that the metrics \( g_{\mu\nu} \) and \( \delta^\mu_{\nu} \) are used to lower or raise indices of tensors on \([q]\) and \([x]\), respectively, the embedding condition (13) can be written in a more general form

\[
\frac{dT^\mu_{\lambda\beta}}{dt} = \frac{d}{dt} \left( \epsilon^\mu_{\nu} \epsilon^\nu_{\lambda} \epsilon^{\beta}_{\gamma} \cdots T^\rho_{\lambda\gamma} \cdots \right), \tag{18}
\]

where the covariant derivative reads

\[
\frac{DT^\mu_{\lambda\beta}}{dt} = v^\lambda \left( \partial_v T^\mu_{\alpha\beta} + \Gamma^\mu_{\alpha\lambda} T^\rho_{\alpha\beta} \cdots \right) - \Gamma^\mu_{\alpha\lambda} T^\rho_{\alpha\beta} \cdots. \tag{19}
\]

Doing the differentiation in the left-hand side of (18) and applying relations (17) we find

\[
v^\mu \partial_v T^\mu_{\nu} = \left( \epsilon^\mu_{\nu} \frac{d}{dt} \epsilon_{\nu} \right) = - \left( \epsilon_{\nu} \frac{d}{dt} \epsilon^\mu_{\nu} \right), \tag{20}
\]

which should hold for any curve in \([q]\) (for any \( v^\mu \)). Thus, we conclude that

\[
\Gamma^\mu_{\nu\lambda} = g^{\lambda\nu} \epsilon_{\nu}(\epsilon_{\lambda}, \epsilon_{\mu}). \tag{21}
\]

Eq. (20) ensures that along any curve \( q^\mu(t) \), the fields \( \epsilon^\mu_{\nu}(q(t)) \) and \( \epsilon^{i\mu}(q(t)) \) are transported parallel, as expressed by the relations \( D \epsilon^{\mu}_{\nu}(q(t))/dt = 0, D \epsilon^{i\mu}(q(t))/dt = 0 \). Applying the covariant derivative \( D/dt \) to the metric (15) we obtain from the chain rule of differentiation \( D g_{\mu\nu}(q(t))/dt = 0 \) for any curve in \([q]\), which ensures the compatibility of the induced connection with the induced metric.

Thus we have succeeded in determining metric and parallel transport in the space \([q]\) by an embedding of all paths in \([q]\) into the space of all paths in the bigger euclidean space. The embedding of the path space implies the embedding of the tangent space, thus determining the induced metric on \([q]\). By imposing the condition that the parallel transport law is compatible with the embedding of the tangent space, the connection in the space \([q]\) is uniquely determined, too.

3. Let us now turn to the analysis of the connec-
Relation 8 can then be written in the form

\[ q \ 

embedding of \ x \ 

into the straight lines in the embedding space \ x. \ 

The torsion induced by the embedding is zero iff

\[ e^i_{\nu,\mu} = e^i_{\mu,\nu}. \] (23)

If this condition is satisfied, the matrix elements in relations (8) are the derivatives of \ M \ functions \ e^i(q) \n
\[ e^i_{\nu}(q) = \partial_\nu e^i(q). \] (24)

In this case, the path embedding \ (8) \ can be achieved by a pointwise embedding of the space \ [q] \ into \ [x]. \ Rela-

\[ 26. \] (25),

i.e., we get the pointwise embedding

\[ x^i = e^i(q). \] (26)

The condition \ (8) \ may be thought as constraints on the velocity \ v. \ The torsion tensor \ (22) \ vanishes when the constraints are \ integrable \ as can be seen from \ (23). \ In this case, the path embedding and the tangent space embedding can be obtained right-away from the space embedding \ (26). \ When the con-
straints are \ non-integrable, \ there is no pointwise embedding of \ [q] \ into \ [x], \ while the path space or the tangent space can still be embedded in the correspond-
ing larger space. The latter is sufficient to specify the metric and connection induced by the embedding. In fact, in this approach the connection appears to be the most general connection compatible with the metric.

The metric tensor \ g_{\mu\nu} \ has \ N(N+1)/2 \ independent components. The torsion tensor \ S_{\mu\nu} \ has \ N^2(N-

\[ 1)/2 \ independent components. To embed a gen-

eral metric space with torsion, the number \ NM, \ being the number of independent embedding coefficients \ e^i_{\mu}, \ should be greater or equal to \ N(N^2+1)/2. \ This leads to the relation between the dimen-

\[ (\dim[q])^2 + 1 \leq 2\dim[x]. \] (27)

4. We are now ready to show that the autoparallel curves are specially favored geometric curves in the space \ [q] \ since our embedding procedure maps them into the straight lines in the embedding space \ [x]. A straight line parallel-transport its tangent vector along itself with respect to a trivial connection \ \[ r^k = 0; \]

\[ \frac{Dv^\mu}{dt} = \dot{v}^\mu = 0. \] (28)

Applying the compatibility condition \ (13) \ to the velocity vector \ T^\mu = v^\mu, \ we conclude that the straight line is the image of a curve satisfying the equation \ \[ \frac{Dv^\mu}{dt} = 0, \] \ which is the autoparallel, as announced above. Indeed, upon substituting \ (8) \ into \ (28) \ and multiplying the result by \ e^j_{\mu} \ we get

\[ (e^\mu, \dot{v}) = \dot{v}^\mu + \left( e^\mu, \frac{d}{dt} e^\nu \right) v^\nu = \frac{Dv^\mu}{dt} = 0. \] (29)

For every path in \ [q], \ the compatibility of the parallel transport law with the embedding yields the condition \ (20). \ Therefore Eq. \ (29) \ turns into the autoparallel Eq. \ (7). In particular, when the con-
straint \ (8) \ is integrable, the torsion vanishes, and Eq. \ (29) \ describes geodesics in the embedded manifold \ (26).

5. It is useful to give a mechanical interpretation of why autoparallels should be favored as particle trajectories. They describe a constrained motion with an acceleration that deviates \ minimally \ from the acceleration of the corresponding unconstrained motion. This property can be formulated mathematically by means of Gauss’ principle of least constraint [2,3].

Consider a Lagrangian system in the space \ [x] \ with a Lagrangian \ \[ L = L(x, \dot{x}). \] At each moment of time, a state of the system can labeled by the pair \ \[ \psi = (x(t), \dot{x}(t)). \] At the state \ \[ \psi, \] \ construct a matrix \ \[ H_{ij} = \frac{\partial^2 L}{\partial \dot{x}^i \partial \dot{x}^j} \] \ called a Hessian of the system. Consider two paths \ \[ x_1(t) \] \ and \ \[ x_2(t) \] \ going through the state \ \[ \psi. \] Gauss’ deviation function (sometimes also called Gauss’ constraint) for two paths at the state \ \[ \psi \] \ reads

\[ G_\psi = \frac{1}{2} (\dot{v}_1^i - \dot{v}_2^i) H_{ij} (\dot{v}_1^j - \dot{v}_2^j). \] (30)

It measures the deviation of two motions from one another at the same system state [2,3]. Let the motion in \ [x] \ be subject to constraints. All paths \ \[ x'(t) \] \ allowed by the constraints and going through a state \ \[ \psi \] \ are called \ conceivable \ motions. A path \ \[ x'(t) \] is
called released motion if it satisfies the Euler-Lagrange equations for the Lagrangian \( L \). Gauss’ principle of least constraint says that the deviation of conceivable motions from a released motion takes a stationary value on the actual motion.

In our case, the released motion is a free motion with zero acceleration \( \ddot{x} = 0 \) and \( H_{ij} = \delta_{ij} \). Accelerations of the conceivable motions are

\[
\ddot{v}^i = e^i_\mu \dot{\nu}^\mu + e^i_\mu \nu^\rho \nu^\rho.
\]  

(31)

Gauss’ deviation function (30) assumes the form

\[
G_\theta = \frac{1}{2} \left[ \dot{v}^\mu + \left( e^\mu_\nu \frac{d}{d\tau} e^\nu_\nu \right) \nu^\nu \right]^2.
\]  

(32)

It is non-negative, \( G_\theta \geq 0 \), for any state \( \psi \) and achieves its absolute minima, \( G_\theta = 0 \), for the acceleration determined by Eq. (29). Thus, the actual motion is realized by autoparallels.

Gauss’ principle of least constraint implies that the geometrical force caused by the constraints, must be minimal for the actual constrained motion. Thus, in the framework of constrained dynamics, autoparallels have the least deviation from the straight lines describing free, unconstrained motions. That is why they can rightfully be called the straightest lines is a consequence of the physical phenomenon inertia. It is hard to understand how a particle should know where to go to make its trajectory the shortest path to a distant point.

Finally, we remark that in the case of integrable constraints, Gauss’ principle of least constraint leads to geodesics, while for non-integrable constraints, geodesics do not have the least deviation from the free, unconstrained motion and do not play any special role in the dynamics.

Another interpretation of the autoparallel Eq. (29) rests on the d’Alembert-Lagrange principle [2,3]. In theoretical mechanics, elements of the tangent space are also called virtual velocities. The embedding condition (10) determines virtual velocities of the constrained motion. Let us denote the Lagrange derivative as \( [L] = d/d\theta \odot \partial L/\partial \nu^i - \partial L/\partial x^i \). The d’Alembert-Lagrange principle asserts that the conceivable motion of a system with the Lagrangian \( L \) is an actual motion if for every moment of time

\[
(T,[L]) = 0,
\]  

(33)

for all virtual velocities of the constrained motion. Taking the free motion Lagrangian \( L = (v,v)/2 \) with the constraint (8) and substituting them into (33) we find that the autoparallel Eq. (29) follows from (33) for an arbitrary virtual velocity \( T^\mu \).

In addition we remark that autoparallels can be obtained from Hölder’s variational principle [2,3] applied to the free action. Let \( \delta x^i \in T[x] \) be a variation vector field. Amongst all variation vector fields we single out those that are virtual velocities of the constrained motion,

\[
\delta x^i = e^i_\mu \delta q^\mu , \quad \delta q^\mu \in T[q].
\]  

(34)

A conceivable path is called a critical point of the action functional if its variation vanishes when restricted on the subspace of virtual velocities of the constrained motion. Hölder’s variational principle suggests that the actual constrained motion is a critical point of the action. Making a variation of the action we find the equation of motion

\[
(\delta x^i,[L]) = 0.
\]  

(35)

Substituting the expression (8) for conceivable velocities and admissible variations (34) in (35), we obtain the autoparallel Eq. (29) for the free Lagrangian, \( [L] = \delta_j \dot{\nu}^j \).

Another remark concerns relativistic motion. The autoparallel motion may also be embedded into a Minkovski space in the same fashion as for a euclidean space. In all formulas given above, the euclidean metric \( \delta_{ij} \) has to be replaced by a corresponding indefinite metric of the Minkovski space. Similarly, in the three variational principles for autoparallels considered in this section, the free motion in the embedding space should be described by the corresponding Lagrangian of a free relativistic particle, while all time derivatives in the equations of motion must be replaced by the derivatives with respect to the proper time defined on the constraint surface.

Finally we point out that the motion of a holonomic system is completely determined by the restriction of the Lagrangian to the constraint surface [2]. Thus, holonomic constrained systems are indistinguishable from ordinary unconstrained Lagrangian systems. This is not true for non-holonomic systems, meaning that the Euler-Lagrange equations for the Lagrangian restricted on the constraint surface do not
coincide with the original equations for the constrained motion. This difficulty prevents us from applying a conventional Hamiltonian formalism to the autoparallel motion, and subjecting it to a canonical quantization. In other words, Dirac’s method of quantizing constrained systems [9] does not apply to non-holonomic systems because their motion is not described by the conventional Lagrange formalism [2]. The problem requires a further study.

Acknowledgements

We are grateful to Drs. G. Barnich, A. Pelster, and A. Tchervyakov for useful discussions. This work was supported in part by the Alexander von Humboldt foundation, by the Deutsche Forschungsgemeinschaft under Project Kl-256/26-1, and by the European Science Foundation, an association of 62 major national funding agencies.

References


Quantum Dirac constraints, Ward identities and path integral in relativistic gauge

A.O. Barvinsky

Theory Department, Lebedev Physics Institute and Lebedev Research Center in Physics, Leninsky Prospect 53, Moscow 117924, Russia

Received 11 December 1997; revised 17 March 1998

Abstract

Quantum Dirac constraints in generic constrained system are solved by directly calculating in the one-loop approximation the gauge field path integral with relativistic gauge fixing procedure. Explicit mechanism of transition from relativistic gauge conditions to unitary gauges, participating in the construction of this solution, is revealed by the method of Ward identities.

1. Introduction

One of the old problems in high energy physics is a path integration in Dirac quantization of constrained systems – the only regular method in problems that go beyond the scope of exactly solvable models. In such systems the Schrodinger equation is supplemented by quantum Dirac constraints on quantum states. Moreover, in parametrized systems with a vanishing Hamiltonian there is not independent Schrodinger equation, and quantum dynamics is encoded in the Dirac constraints along with gauge invariance properties. For such problems the fundamental dynamical equations are not of a manifestly evolution type and, therefore, the path integral applications are much less straightforward.

One of the first formulations of the path integral as a solution of the non-evolutionary quantum Dirac constraints belongs to Leutwyler [1] who proposed a two-point solution of the Wheeler-DeWitt equations in the form of a naive functional integral. The path integral with unitary gauge fixing procedure and exhaustive set of boundary conditions was later proposed in [2]. Then this canonical path integral was converted to the spacetime covariant form of the functional integral over Lagrangian variables and ghost fields [3–5].

Of course, these results incorporate a well-known statement on equivalence of the canonical and covariant quantizations pioneered in [6]. In contrast with this, the works [2–5] were focused on the nontrivial boundary conditions in spacetime. Their correct treatment leads to the proof that this path integral solves the quantum Dirac constraints. However, this proof [2,3,7,5] has a formal nonperturbative nature and does not even allow one to fix the operators of quantum constraints which in a rather uncontrollable way depend on the calculational method for a path integral [7]. Thus, no check of the solution to quantum constraints was thus far given.
by direct calculations of the path integral. The goal of this paper is to perform such a check in the one-loop approximation of semiclassical expansion. The calculations are based on the reduction methods for one-loop functional determinants [8] extended here to the case of special boundary conditions characteristic of field theory. They can be distinguished by context in which they appear. For Einstein theory, they are subject to canonical commutation relations of time, i.e., of functional matrices with respect to the functional determinants of operators on the space of functions of time, i.e., of functional matrices with covariant condensed indices involving time, and their contraction also implies time integration. They can be distinguished by context in which they appear. For Einstein theory, they are subject to canonical commutation relations of time, i.e., of functional matrices with respect to the functional determinants of operators on the space of functions of time, i.e., of functional matrices with covariant condensed indices involving time, and their contraction also implies time integration.

2. Dirac quantization of constrained systems

A theory with the Lagrangian action \( S[ g ] \) invariant under local gauge transformations of fields \( g = g^a = (q^a(t), N^\mu(t)) \) with generators \( R^a_\mu \)

\[
S[ g ] = \int_{t_1}^{t_2} dt L(q, \dot{q}, N), \quad R^a_\mu \frac{\delta S[ g ]}{\delta g^a} = 0
\]  

(2.1)

in the canonical formalism is parametrized by phase space variables \((q^a, p_a)\) and Lagrange multipliers \( N^\mu = N^\mu(t) \) for the first class constraints

\[
T^a_\mu(q, p) = U^a_\mu T^a_\mu, \quad \text{with structure functions } U^a_\mu.
\]

Dirac quantization consists in promoting phase-space variables and constraints to the operator level \((q^a, p_a) \rightarrow (\hat{q}^a, \hat{p}_a)\) and selecting the physical states \(|\Psi\rangle\) in the representation space of \((\hat{q}^a, \hat{p}_a)\) by the equation \(\hat{T}^a_\mu |\Psi\rangle = 0\) \([12, 11, 10, 13]\). Operators \((\hat{q}^a, \hat{p}_a)\) are subject to canonical commutation relations of the form

\[
[\hat{q}^a, \hat{p}_b] = i\hbar \delta^a_b
\]

and operators \(\hat{T}^a_\mu\) should satisfy the Poisson bracket algebra

\[
[\hat{T}^a_\mu, \hat{T}^b_\nu] = i\hbar \delta^a_b T^b_\mu.
\]

It should be noted that the Dirac brackets are not the commutators of the operators \(\hat{T}^a_\mu\).

The symbols here and in (2.4) denote the determinants with respect to canonical condensed indices in contrast with \(\det \) – the functional determinants of operators on the space of functions of time, i.e., of functional matrices with covariant condensed indices.
to a formal proof of quantum Dirac constraints [4,7]. This nonperturbative derivation is, however, purely formal and in an uncontrollable way depends on the skeletonization of the path integral [7].

The semiclassical representations of $K(q,q')$ can be obtained by directly solving quantum constraints or calculating the path integral (2.2). The one-loop solution $K(q,q') = P(q,q') e^{iW(q,q')}$ (2.5) contains the Hamilton-Jacobi function $S(q,q')$ – the action (2.1) at the classical extremal which joins the points $q$ and $q'$ and the preexponential factor $P(q,q')$ [9,10,14]. The latter was obtained in [9,10] as a special generalization of the Pauli-Van Vleck-Morette formula [15] in terms of the degenerate Van-Vleck matrix [9,10].

\begin{equation}
S_{ik} = \frac{\partial^2 S(q,q')}{\partial q^i \partial q'^k}.
\end{equation}

(2.6)

\[ \nabla^\mu S_{ik} = 0, \quad S_{ik} \nabla^\mu = 0, \quad \nabla^\mu = \frac{\partial T^\mu}{\partial p_i}, \quad \nabla^\nu = \frac{\partial T^\nu}{\partial p^k}, \quad \nabla^\mu \nabla^\nu = D_{ik}.
\]

(2.7) This generalization is equivalent to the gauge-fixing procedure. It consists in adding to (2.6) the “gauge-breaking” term quadratic in “gauge conditions” $X_i^\mu$ and $X_i^\nu$ – two sets of arbitrary covectors at the points $q$ and $q'$ which produce invertible “Faddeev-Popov operators” $J_i^\mu$ and $J_i^\nu$.

\begin{equation}
D_{ik} = S_{ik} + X_i^\mu C_{ik} X_k^\nu.
\end{equation}

(2.9)

\begin{equation}
J_i^\mu = X_i^\mu C_{ik} X_k^\nu, \quad J \equiv \det J_i^\mu, \quad J_i^\nu = X_i^\mu C_{ik} X_k^\nu, \quad J' = \det J_i^\nu.
\end{equation}

(2.10)

($C_{ik}$ is some invertible gauge-fixing matrix). Then the one-loop prefactor reads

\begin{equation}
P(q,q') = \left[ \frac{\det D_{ij}}{JF \det C_{ik}} \right]^{1/2}.
\end{equation}

(2.11) This solution is a direct analogue of the one-loop effective action in gauge theory – the gauge fields contribute $\det D_{ik}$ partly compensated by that of ghosts $J\beta$.

The covectors $(X_i^\mu, X_i^\nu)$ play the role of linearized gauge conditions because, as shown in [9,10], the quantum Hamiltonian reduction of $K(q,q')$ leads to the unitary evolution operator in the physical sector defined by the unitary gauge conditions related to $X_i^\mu$

\begin{equation}
X_i^\mu(q,t) = 0, \quad X_i^\nu = \partial X^\mu / \partial q^i.
\end{equation}

(2.12) The unitary gauge conditions are imposed only on phase space variables $(q,p)$ (in this case only on coordinates $q$) and therefore break manifest covariance.

On the contrary, the path integral (2.2) in relativistic gauge is potentially a spacetime covariant object. Its Feynman diagrammatic technique was built in [4] with a special emphasis on boundary conditions at $t\pm$, and in the one-loop order it also has the form (2.5) with the same principal Hamilton function – the action at the solution $q = q'(t \mid q_\pm q_-)$ of the classical boundary value problem in the gauge

\begin{equation}
q(t_\pm) = q _\pm,
\end{equation}

(2.13) and the one-loop preexponential factor

\begin{equation}
P(q_+,q_-) = \left[ \frac{\det F_{ab}}{\det a_{ab}} \right]^{-1/2} \frac{\det Q^\nu_{\mu}}{\det a^\nu_{\mu}} \left|_{g = g(t \mid q_+,q_-)} \right.
\end{equation}

(2.14) Here $F_{ab}$ is the gauge field operator

\begin{equation}
F_{ab} = S_{ab} - \chi^\mu_{ab} c_{\mu\nu} X_{b}. \quad S_{ab} = \frac{\delta^2 S[\chi]}{\delta g^a \delta g^b}, \quad \chi^\mu_{ab} = \frac{\delta \chi^\mu}{\delta g^a},
\end{equation}

(2.15) with the functional matrix $\chi^\mu_{ab}$ of linearized gauge conditions which is a first-order differential operator $\chi^\mu_{ab} = \chi^\mu_{ab} (d/dt) \delta(t-t')$.

The determinants in (2.14) are calculated on functional spaces defined by boundary conditions for gauge and ghost operators. In relativistic gauges (2.4) these operators $F_{ab} = (-a_{ab} d^2/dt^2 + \ldots) \delta(t-t')$ and $Q^\nu_{\mu} = (-a^\nu_{\mu} d^2/dt^2 + \ldots)$ are of second order in time derivatives with nondegenerate matrix-
ces $a_{ab} = \delta^2 L_{\text{gf}} / \delta g^a \delta g^b$ (the Hessian of gauge-fixed Lagrangian) and (2.4). Their boundary conditions were derived in [4] from the boundary conditions on integration variables in the path integral. For $(n + m) \times (n + m)$-matrix of the gauge field propagator $G^{ab} = G^{ab}(t,t')$ they form a combined set of $n$ Dirichlet conditions $X^{\mu} (d/dt) G^{ab}(t,t') = 0$. This leads to the relation $R^\mu_{ab} = -Q_a c_{ab} \chi^\mu$, which implies the following Ward identity

$$
\rightarrow \beta \chi^\mu_{ab}(t,t') = -Q_a c_{ab} \chi^\mu_{ab}(t,t').
$$

Here the arrows show the direction in which the differential operators $\chi^\mu_{ab}(d/dt)$ and $R^\mu_{ab}(d/dt)$ are acting on the arguments of Green’s functions.

These Ward identities lead to gauge independence of the one-loop gauge-fixing prefactor (2.14) provided we consistently fix the definition of functional determinants $F_{ab}$ and $Q^\mu$. Problem is that unique Green’s functions do not yet uniquely fix them. Variational equations for determinants involve the functional composition of Green’s functions with variations of the corresponding operators: $G^{\mu \nu} F_{\mu \nu}$ and $Q^{-1}_{\nu} \delta Q^\mu$. However, kernels of propagators are not smooth functions and their irregularity enhances when they are acted upon by the derivatives of $\delta F_{ab}(d/dt)$ and $\delta Q^\mu(d/dt)$. Therefore, integrations by parts can lead to surface terms nonvanishing even despite correct boundary conditions. We assume that $\delta F_{ab}(d/dt)$ and $\delta Q^\mu(d/dt)$ are acting in two different ways

$$
\delta \ln \det F_{ab} = G^{\mu \nu} J_{ab}^\mu \nu, 
$$

$$
\delta \ln \det Q^\mu = Q_{\nu}^{-1} \delta Q^\mu_{\nu}.
$$

In the covariant context the analogue of (2.7) is the on-shell degeneracy of the Hessian matrix – a corollary of (2.1)

$$
R^\mu_{ab} = \delta^2 S \left[ g \right] \frac{\delta \delta S}{\delta g^a \delta g^b} \bigg|_{\delta g = 0} = 0.
$$

It follows from the Ward identities for the gauge-fixed Van Vleck Green’s function, Eq. (3.4) implies a symmetric action of $\delta F_{ab}(d/dt)$ on both arguments of $G^{\mu \nu}(t,t')$ in the sense that $L^{(2)}_{\text{gf}} = (1/2) \varphi'(t) F_{ab}(d/dt) \varphi'(t)$ represents the symmetric part of the gauge-fixed Lagrangian in field perturbations $\varphi'$ (and $\delta F_{ab}(d/dt)$ obviously implies arbitrary variation of its kernel).
With these conventions the gauge variation of gauge and ghost determinants in (2.14)

\[
\delta \chi \ln \text{Det} F_{ab} = -2 \epsilon_{\alpha \beta} \chi^b_a G^{ba} \delta \chi^a_b ,
\]

(3.6)

\[
\delta \chi \ln \text{Det} Q_{\mu}^{\alpha b} = Q_{\mu}^{\alpha b} \delta \chi^a_b ,
\]

(3.7)
cancel out in virtue of (3.3). This proves gauge independence of the expression (2.14) along with fixing the prescription for the variational definition (3.4), (3.5) of its functional determinants.

4. Reduction algorithms for functional determinants

Equality of expressions (2.11) and (2.14) follows from the reduction algorithms for functional determi-

nants which reduce their dimensionality from the functional dimensionality of Det’s in (2.14) to that of det’s in (2.11) [8]. The simplest example of such algorithms is the expression for a purely divergent det’s in (2.11) 8 . The simplest example of such

4. Reduction algorithms for functional determinants

Equality of expressions (2.11) and (2.14) follows from the reduction algorithms for functional determi-

nants which reduce their dimensionality from the functional dimensionality of Det’s in (2.14) to that of det’s in (2.11) [8]. The simplest example of such algorithms is the expression for a purely divergent det’s in (2.11) 8 . The simplest example of such

4. Reduction algorithms for functional determinants

Equality of expressions (2.11) and (2.14) follows from the reduction algorithms for functional determi-

nants which reduce their dimensionality from the functional dimensionality of Det’s in (2.14) to that of det’s in (2.11) [8]. The simplest example of such algorithms is the expression for a purely divergent det’s in (2.11) 8 . The simplest example of such

4. Reduction algorithms for functional determinants

Equality of expressions (2.11) and (2.14) follows from the reduction algorithms for functional determi-

nants which reduce their dimensionality from the functional dimensionality of Det’s in (2.14) to that of det’s in (2.11) [8]. The simplest example of such algorithms is the expression for a purely divergent det’s in (2.11) 8 . The simplest example of such

4. Reduction algorithms for functional determinants

Equality of expressions (2.11) and (2.14) follows from the reduction algorithms for functional determi-

nants which reduce their dimensionality from the functional dimensionality of Det’s in (2.14) to that of det’s in (2.11) [8]. The simplest example of such algorithms is the expression for a purely divergent det’s in (2.11) 8 . The simplest example of such

4. Reduction algorithms for functional determinants

Equality of expressions (2.11) and (2.14) follows from the reduction algorithms for functional determi-

nants which reduce their dimensionality from the functional dimensionality of Det’s in (2.14) to that of det’s in (2.11) [8]. The simplest example of such algorithms is the expression for a purely divergent det’s in (2.11) 8 . The simplest example of such

4. Reduction algorithms for functional determinants

Equality of expressions (2.11) and (2.14) follows from the reduction algorithms for functional determi-

nants which reduce their dimensionality from the functional dimensionality of Det’s in (2.14) to that of det’s in (2.11) [8]. The simplest example of such algorithms is the expression for a purely divergent det’s in (2.11) 8 . The simplest example of such
terms at $t = t_\pm$ which form the total variation that can be integrated to give [17]:

$$
\left( \frac{\det F_{ab}}{\det a_{ab}} \right)^{-1/2} = \text{const} \left( \frac{\det \left( S_{ik} + X_\mu^\nu C_{\mu
u} X_\nu^\nu \right)}{\det C_{\mu
u}} \right)^{1/2}.
$$

(4.8)

Here the matrices $(X_\mu^\nu, X_\nu^\nu, C_{\mu
u})$ read in terms of basis functions of the above type

$$
X_\mu^\nu = (W_{\mu,+})(t_+) \Delta_{1/2}^{-1} u_\mu^\nu(t_-),
$$

(4.9)

$$
X_\nu^\nu = u_\nu^\nu(t_+) \Delta_{1/2}^{-1} (W_{\nu,-})(t_-),
$$

(4.10)

$$
C_{\mu\nu} = u_\mu^\nu(t_+) \Delta_{1/2}^{-1} u_\nu^\nu(t_-),
$$

(4.11)

and $S_{ik}$ arising here as $S_{ik} = -(W_{\mu,+})(t_+)$ $\times$ $\Delta_{1/2}^{-1} (W_{\nu,-})(t_-)$ coincides with the Van Vleck matrix (2.6) [17]. This equation generalizes the results of [8] by extending the path-integral derivation of the Pauli-Van Vleck-Morette formula [15] to gauge theories. Its comparison with (2.9),(2.11) shows that the quantities (4.9)–(4.11) describe a special unitary gauge induced by the relativistic gauge fixing procedure.

The functional integration of Eq. (3.5) for $\det Q_\nu^{\mu}$ repeats the calculations of [8] modified by the asymmetry of the ghost operator. This asymmetry results in a double set of right and left basis functions which give rise to the representation of $Q_\nu^{\mu}$ similar to (4.7). Functional integration yields the analogue of the Pauli-Van Vleck formula with a special Van Vleck matrix [17]. In view of Ward identity (3.3) special combinations of ghost in this matrix can be related to the unitary gauge conditions (4.9), (4.10) built in terms of the basis functions of the gauge operator. The final reduction algorithm then reads

$$
\frac{\det Q_\nu^{\mu}}{\det a_\nu^{\mu}} = \text{const} \left( \det J_\mu^{\nu} \det J_\nu^{\mu} \right)^{-1/2},
$$

(4.12)

where the matrices $J_\mu^{\nu}$ and $J_\nu^{\mu}$ coincide with unitary Faddeev-Popov operators (2.10) in unitary gauges (4.9), (4.10). The substitution of (4.8) and (4.12) to (2.14) accomplishes the proof for the needed equality of (2.11) and (2.14).

5. Conclusions

The virtue of equivalence of (2.11) and (2.14) is that it explicitly establishes the mechanism of transition from relativistic to unitary gauge conditions. The relation between them is nonlocal in time – the matrices of unitary gauge-fixing procedure express in terms of basis functions of the gauge field operator and nonlocally depend on its relativistic gauge. Such a transition proves intrinsic unitarity of a manifestly covariant quantization by a Lagrangian path integral without singularities inherent to the usual $\epsilon$-procedure of formally identical transformations in the path integral [6].

In physical applications, Feynman diagrammatic expansion of the path integral as a means of solving noncovariant Dirac constraints [19] is of crucial importance due to spacetime covariance of their solution that can be attained by a suitable choice of relativistic gauge conditions. Important implications of this technique belong to quantum cosmology of the early universe [18] where correct predictions can be achieved only within spacetime covariant approach to loop effects.

The aspects of gauge independence are also important in spacetimes with boundaries or nontrivial time foliations. There exists a long list of gauge-dependent results for a formally gauge independent quantity – the one-loop effective action [20]. Exhaustive explanation of these discrepancies can be expected on the basis of Ward identities of the above type.

Finally, the physics of wormholes in Euclidean quantum gravity [21] also belongs to the scope of our result. The predictions of this theory are based, in particular, on the existence of a negative mode on the wormhole instanton [22] – a formal extrapolation of the mechanism applicable only to non-gravitational systems [23]. Thus, these predictions should be revised from the viewpoint of the Wheeler-DeWitt equation [24,22]. This negative mode belongs to the nondynamical conformal sector, and in the
Lorentzian theory its contribution is cancelled by ghost fields in relativistic gauges. Therefore, one should expect a similar cancellation in Euclidean theory. Other problems related to the indefiniteness of the Euclidean gravitational action include the lack of strong ellipticity of the boundary value problem (2.16), (2.17) [25]. The proposed technique can be regarded as a direct avenue towards the resolution of these issues which are currently under study.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research under grants 96-02-16287 and 96-02-16295 and the European Community Grant INTAS-93-493-ext. Partly this work has been made possible also due to the support by the Russian Research Project "Cosmomicrophysics".

References

Hamiltonian flow equations for a Dirac particle in an external potential

A.B. Bylev a, H.J. Pirner b

a Institute of Physics, St-Petersburg State University, Germany
b Institute of Theoretical Physics, Heidelberg University, Germany

Received 22 December 1997; revised 6 March 1998
Editor: P.V. Landshoff

Abstract

We derive and solve the Hamiltonian flow equations for a Dirac particle in an external static potential. The method shows a general procedure for the set up of continuous unitary transformations to reduce the Hamiltonian to a quasidiagonal form.

The well known Foldy-Wouthuysen FW transformation [1] is a unitary transformation that separates big and small components in the Dirac equation for a particle interacting with an external electromagnetic field. If the external field is stationary then the FW transformation reduces to a unitary transformation of the Hamiltonian $H$ of the Dirac particle

$$H' = e^{iS}(\alpha(p - eA(x)) + \beta m + eA_\beta(x))e^{-iS}$$

(1)

where $S$ can be expanded in a sequence of hermitian operators in powers of $1/m$. Recently a novel method to diagonalize Hamiltonians via continuous unitary transformations has been proposed [2], where the Hamiltonian $H$ and the $S$ are considered to be functions of a continuous variable, the flow parameter $l$. This new method has been applied to several topics like the $n$– orbital model [2], superconductivity and the electron phonon interaction [3,4]. We want to apply it here to the Dirac particle in an external field in order to explore its possibilities with a well known problem. Thereby we can gain experience in solving problems like the elimination of negative energy states in near light cone QCD [5], where the solution has yet to be discovered.

Support by DFG Grant 436 RUS 113/205/1 on Russian German Cooperation.

0370-2693/98/$19.00 © 1998 Elsevier Science B.V. All rights reserved.
PII: S0370-2693(98)00417-1
For infinitesimal steps \( \exp(\eta(l)dl) \) a continuous unitary transformation of the Hamiltonian \( H \) leads to the so-called Hamiltonian flow equation

\[
\frac{dH(l)}{dl} = [\eta(l), H(l)] \tag{2}
\]

with initial condition \( H(0) = H \). The transformation matrix \( \eta(l) \) is antihermitian and related to \( S \) as follows:

\[
\eta(l) = \left( \frac{d}{dl} e^{iS(l)} \right) e^{-iS(l)}. \tag{3}
\]

The parameter \( l \) has dimension of energy \( \gamma^2 \) and characterizes the amount of diagonalization which has been achieved. In the energy representation the matrix elements of the Hamiltonian \( H \) will be grouped in a band of width \( \Delta E = \frac{\gamma}{\sqrt{l}} \) around the diagonal, if \( \eta(l) \) is chosen in the correct way. The result of the transformation depends critically on the choice of \( \eta(l) \) as a function of \( H(l) \). In Ref. [2] the following expression has been proposed:

\[
\eta(l) = [H_D(l), H(l)] \tag{3}
\]

where \( H_D(l) \) is the diagonal part of \( H(l) \). Indeed such a choice of \( \eta(l) \) leads to a quasidiagonalization of \( H(l) \) when \( l \to \infty \). From Eqs. (2) and (3) we obtain that

\[
\frac{d}{dl} \text{tr}H_D^2(l) = -\frac{d}{dl} \text{tr}H_D^2(l) = 2\text{tr}([H(l), H_D(l)] [H_D(l), H(l)]) \geq 0.
\]

Thus the sum of the absolute squares of the nondiagonal elements decreases, whereas the sum of the diagonal matrix elements squared increases with \( l \). However, this line of reasoning only holds if one can solve Eqs. (2), (3) exactly. For approximate solutions of Eqs. (2) and (3) divergencies can arise [2].

In this report we want to present a continuous transformation of the Dirac Hamiltonian via flow Eqs. (2) the derivation of which is guided by the FW transformation. The basic idea of the FW transformation is a decomposition of the Hamiltonian into two noninteracting parts corresponding to positive and negative energies. Therefore, the transformed Hamiltonian must commute with the \( \beta \) matrix and we choose as ansatz \( \eta(l) \) in the form

\[
\eta(l) = [\beta m, H(l)]. \tag{4}
\]

The Hamiltonian \( H(l) \) can be presented as a sum of an even operator \( \varepsilon(l) \) and odd operator \( \sigma(l) \), where the even or oddness is defined by the commutation relations of the respective operators with the \( \beta \) matrix.

\[
H(l) = \sigma(l) + \varepsilon(l),
\]

\[
\varepsilon(l) \beta = + \beta \varepsilon(l),
\]

\[
\sigma(l) \beta = - \beta \sigma(l).
\]

The initial conditions for the evolution equation

\[
\frac{dH(l)}{dl} = [\eta(l), H(l)]
\]

are obtained by using the plus commutator \([\alpha, \beta]_+ = 0\).

\[
\varepsilon(0) = \beta m + eA_0,
\]

\[
\sigma(0) = \alpha (p - eA).
\]

Then our ansatz for the transformation matrix \( \eta(l) \) becomes

\[
\eta(l) = 2m \beta \sigma(l).
\]
The flow equation can be split up into the following system of two equations:

$$\frac{d\mathcal{E}(l)}{dl} = 4m\beta\mathcal{E}^2(l),$$

$$\frac{d\mathcal{E}(l)}{dl} = 2m\beta [\mathcal{E}(l),\mathcal{E}(l)].$$

As we will show now, the system of Eqs. (5), (6) can be solved perturbatively in $1/m$. It is convenient to introduce a dimensionless flow parameter $\lambda = l \cdot m^2$. Since $\mathcal{E}(0) = \beta m + eA_0$ the expansion of $\mathcal{E}(\lambda)/m$ in a series in $1/m$ contains terms starting with the zeroth order term

$$\frac{1}{m}\mathcal{E}(\lambda) = \mathcal{E}_0(\lambda) + \frac{1}{m^2}\mathcal{E}_2(\lambda) + \ldots$$

whereas the expansion of $\mathcal{E}(\lambda)/m$ starts with the first order

$$\frac{1}{m}\mathcal{E}(\lambda) = \mathcal{E}_1(\lambda) + \frac{1}{m^2}\mathcal{E}_2(\lambda) + \ldots.$$

Due to Eqs. (7) and (8) we have in $n$-th order:

$$\frac{d}{d\lambda}\mathcal{E}_n(\lambda) = 4\beta \sum_{k=1}^{n-1} \mathcal{E}_k(\lambda)\mathcal{E}_{n-k}(\lambda),$$

$$\frac{d}{d\lambda}\mathcal{E}_n(\lambda) = -4\mathcal{E}_n(\lambda) + 2\beta \sum_{k=1}^{n-1} \left[\mathcal{E}_k(\lambda),\mathcal{E}_{n-k}(\lambda)\right].$$

The solution of these equations are

$$\mathcal{E}_n(\lambda) = \mathcal{E}_n(0) + 4\beta \int_0^\lambda d\lambda' \left( \sum_{k=1}^{n-1} \mathcal{E}_k(\lambda')\mathcal{E}_{n-k}(\lambda') \right),$$

$$\mathcal{E}_n(\lambda) = \mathcal{E}_n(0) e^{-4\lambda} + 2\beta e^{-4\lambda} \int_0^\lambda d\lambda' \left( \sum_{k=1}^{n-1} \left[ e^{4\lambda} \mathcal{E}_k(\lambda'),\mathcal{E}_{n-k}(\lambda') \right] \right)$$

with initial conditions

$$\mathcal{E}_0(0) = \beta, \mathcal{E}_1(0) = eA_0(x), \mathcal{E}_n(0) = 0 \text{ if } n \geq 2,$$

$$\mathcal{E}_1(0) = \alpha (p - A(x)), \mathcal{E}_n(0) = 0 \text{ if } n \geq 2.$$

One sees that $\mathcal{E}_n(\lambda)$ exponentially goes to zero when $\lambda \to \infty$. By virtue of this behavior integrals in Eq. (11) do not diverge and $\mathcal{E}_n(\lambda)$ is finite when $\lambda \to \infty$.

For the first four orders we have

$$\mathcal{E}_0(\infty) = \mathcal{E}_0(0) = \beta,$$

$$\mathcal{E}_1(\infty) = \mathcal{E}_1(0) = eA_0(x),$$

$$\mathcal{E}_2(\infty) = \frac{1}{2}\beta\mathcal{E}_1^2(0) = \frac{1}{2}\beta \left( (p - eA(x))^2 - eA_0B(x) \right),$$

$$\mathcal{E}_3(\infty) = \frac{1}{8} \left[ [\mathcal{E}_1(0),\mathcal{E}_1(0)],\mathcal{E}_1(0) \right]$$

$$= -\frac{ie}{8} \alpha (\nabla \times E(x)) - \frac{e}{4} \alpha (E(x) \times (p - eA(x))) - \frac{e}{8} (\nabla E(x)).$$
These terms reproduce well known Foldy-Wouthuysen Hamiltonian [1]. The formulae (11), (12) let us simply get expressions in any order in \(1/m\).

Similarly to the renormalization group equations the Hamiltonian flow equations approach the correct solution in infinitesimal steps. We have derived in this new method a solution of the flow Eqs. (2) and (3) for the problem of a Dirac particle in an external potential. This solution has a more general validity. We have shown how to choose the diagonalizing matrix \(\eta\), if the Hamiltonian can be represented as a sum of even and odd operators under commutation with some operator \(D\). Choosing the operator \(\eta(i)\) in the form \([D, H(i)]\) allows us to transform the Hamiltonian \(H\) to a form which commutes with \(D\) whereas the anticommuting part goes to zero when the flow parameter goes to infinity.

The relevance of this approach extends to the interesting problem of quarks with positive and negative light cone energies in near light cone QCD [5]. The SU(3) Hamiltonian in near light front coordinates is formulated in the following way:

\[
\mathcal{H} = -\frac{i}{\eta} \psi_\perp (\partial_+ - ig \alpha_\perp) \psi_\perp - \frac{1}{\eta} \psi_\perp^\dagger \mathbf{\alpha}_\perp (\mathbf{\nabla}_\perp - ig A_\perp') \psi + \frac{1}{\eta} \psi_\perp^\dagger \beta \psi \\
+ \text{tr} \left[ \partial_+ A'_+ - \partial_- A'_- - ig \left[ A'_+, A'_- \right] \right]^2 + \frac{1}{\eta} \text{tr} \left[ \Pi'_+ - \left( \partial_+ A'_- - ig \left[ a'_-, A'_+ \right] \right) \right]^2 \\
+ \frac{1}{\eta^2} \left[ \frac{1}{L} \epsilon_{\perp} - \mathbf{\nabla}_\perp a_\perp \right] + \frac{1}{2L^2} \sum_{\epsilon_{\perp}} p_{\perp}^a (x_{\perp}^+) p_{\perp}^a (x_{\perp}^-) \\
+ \frac{1}{L^2} \int_0^L dx_\perp \int_0^L dy_\perp \sum_{p, q, n} \frac{G_{\perp, q}^p (x_{\perp}^+, z^-) G_{\perp, p}^q (x_{\perp}^+, y^-)}{L} \frac{2\pi n}{L} + g \left( a_{\perp, q} (x_{\perp}^-) - a_{\perp, p} (x_{\perp}^+) \right) \left| e^{\pm i \pi (z^- - y^-) / L} \right|^2
\]

with \(p, q = 1, 2, 3\).

It must be emphasized that the negative energy states \(\psi_\perp = \frac{1}{2}(1 - \alpha^3)\psi\) are dynamical fields in this formulation, whereas on the exact light–front they are constrained fields. In the transverse gluon fields, \(A_{\perp}^a (i = 1, 2 \ldots 8)\), the neutral, two-dimensional longitudinal parts have been eliminated. The conjugate chromoelectric fields are denoted with \(\Pi_{\perp, a}^a\). The near light cone Hamiltonian contains zero mode gluon fields \(a_\perp\) and their canonical momenta \(p_\perp\). Neither of these fields depend on \(x_-\). All the fields are color diagonal. The operator \(G_{\perp, a}^a\) is the transverse part of the Gauss law, the neutral chromo-electric fields \(e_{\perp, a}\), and the neutral charge \(Q_{\perp, a}\) are defined as in Ref. [5]. The parameter \(\eta\) governs the nearness of the theory to the exact light cone.

In principle a continuous unitary transformation similar to the one outlined above can eliminate the negative fermion energy states and generate an effective Hamiltonian for the quarks with positive light cone energies only. A new feature in light cone QCD, however, is the presence of the zero mode gluon fields \(a_\perp\) which couple only to the negative energy states. This is a consequence of the axial gauge with periodic boundary conditions which necessitate the introduction of the \(a_\perp\) fields. A naive elimination of the negative energy states leads to a coupling of the negative energy fermions to the background zero mode fluctuations. Therefore the task is to solve the mixing of positive and negative energy states first before decoupling the new eigenstates in a manner similar to the one outlined in the paper.

Acknowledgements

We thank Andreas Mielke and Evgeni Prokhvatilov for useful discussions. A.B. thanks the Deutsche Forschungsgemeinschaft for the support of his visit to the Institute of Theoretical Physics of Heidelberg University.
References

Thermodynamics of gauge-invariant $U(1)$ vortices from lattice Monte Carlo simulations

K. Kajantie $^{a,b,1}$, M. Karjalainen $^{b,2}$, M. Laine $^{a,b,3}$, J. Peisa $^{c,4}$, A. Rajantie $^{d,5}$

$^a$ Theory Division, CERN, CH-1211 Geneva 23, Switzerland
$^b$ Department of Physics, P.O. Box 9, 00014 University of Helsinki, Finland
$^c$ Department of Mathematical Sciences, University of Liverpool, Liverpool L69 3BX, UK
$^d$ Helsinki Institute of Physics, P.O. Box 9, 00014 University of Helsinki, Finland

Received 19 March 1998
Editor: R. Gatto

Abstract

We study non-perturbatively and from first principles the thermodynamics of vortices in 3d $U(1)$ gauge + Higgs theory, or the Ginzburg-Landau model, which has frequently been used as a model for cosmological topological defect formation. We discretize the system and introduce a gauge-invariant definition of a vortex passing through a loop on the lattice. We then study with Monte Carlo simulations the total vortex density, extract the physically meaningful part thereof, and demonstrate that it has a well-defined continuum limit. The total vortex density behaves as a pseudo order parameter, having a discontinuity in the regime of first order transitions and behaving continuously in the regime of second order transitions. Finally, we discuss further gauge-invariant observables to be measured. © 1998 Elsevier Science B.V. All rights reserved.

PACS: 98.80.Cq; 11.27. + d; 11.15.Ha; 67.40.Vs; 74.60.-w

1. Introduction

Vortices play a significant role from the low temperatures of liquid crystals [1], superfluids [2] and high-$T_c$ superconductors [3], to the relativistic temperatures of the Early Universe [4]. In low temperature systems, vortices can be directly observed [1,5]; in cosmology, one has studied their effects on the inhomogeneities leading to structure formation [6]. Consequently, vortices have been a subject of immense interest during the last few years. Nevertheless, some important questions, related in particular to non-perturbative studies of vortices in gauge theories, remain poorly understood.

Among the most fundamental principles of Nature appear to be gauge invariance and spontaneous symmetry breaking, and already the simplest theory with these properties, the locally $U(1)$ symmetric gauge + scalar quantum field theory or the Ginzburg-Landau (GL) model, contains vortices. The GL model does describe real physics in liquid crystals and superconductors [7], while in cosmology it is to be viewed as
a simple toy model. The phase structure of the GL model is non-trivial: in the type I regime, there is a first order transition [8], whereas in the type II regime, the transition is assumed to be of the second order [9–12].

One of the mentioned open questions arises immediately when one realizes that the type II regime of the GL model is completely non-perturbative: perturbation theory does not describe the transition at all [13]. The only known systematic and controllable method for studying this regime are lattice Monte Carlo simulations [13–18]. Yet as to date, to our knowledge, the vortex density has not been studied in detail on the lattice even in thermodynamical equilibrium. The purpose of this paper is (a) to provide a gauge-invariant formulation for studying vortices on the lattice, (b) to measure the vortex density both in type I and type II regimes, and (c) to extrapolate the results to the infinite volume and continuum limits. The length distribution of vortices will be studied in a future publication [19]. In a U(1) scalar field theory without gauge symmetry, the thermodynamics of vortices has previously been addressed in [20].

Let us stress that considering the thermodynamics of vortices is certainly only a starting point. Ultimately one is interested in the real-time scaling properties of vortex networks created in a non-equilibrium situation see, e.g., [21–23]. However, the thermodynamical equilibrium situation provides the initial conditions for such non-equilibrium processes. Furthermore, it is clear that non-equilibrium physics cannot be understood in quantitative detail before the equilibrium limit is under control.

2. The theory in the continuum and on the lattice

Let us start by defining the theory. The continuum theory is defined by the functional integral

\[ Z = \int \mathcal{D}A \mathcal{D}\phi \exp \left[ -S(A, \phi) \right] , \]

\[ S = \int d^4x \left[ \frac{1}{4} F_{ij}^2 + |D_i \phi|^2 + m_i^2 \phi \cdot \phi + \lambda_i (\phi \cdot \phi)^2 \right] , \]  

where \( F_{ij} = \partial_i A_j - \partial_j A_i \) and \( D_i = \partial_i + ie \phi A_i \). The theory is invariant under the gauge transformations

\[ \phi(x) \to e^{i\theta(x)} \phi(x) , \]

\[ A_i(x) \to A_i(x) - \partial_i \theta(x) / e . \]

Writing \( \phi(x) = \psi(x) \exp(i \gamma(x)) \), the first of these can be rewritten as \( \gamma(x) \to [\gamma(x) + \theta(x)]_\mu \), where \([X]_\mu = X + 2 \pi n \) such that \([X]_\mu \in (-\pi, \pi] \) and we have chosen to represent the phase of \( \phi \) by a number in this interval. The theory in Eq. (2) is parameterized by the scale \( e_3^2 \) and by the two dimensionless ratios

\[ y = \frac{m_i^2(e_3^2)}{e_3^4} , \quad x = \frac{\lambda_i}{e_3^2} , \]  

where \( m_i^2(\mu) \) is the mass parameter in the \( \overline{\text{MS}} \) dimensional regularization scheme in \( 3 - 2 \epsilon \) dimensions. Expressions for \( x \) and \( y \) in terms of the original physical parameters of both 4d high temperature scalar + fermion electrodynamics and 3d low-\( T_c \) superconductivity, have been discussed in [18] where we refer to for more details. Here we just study the theory as a function of \( x, y \). The phase diagram is shown in Fig. 1.

As is well known, the classical counterpart of the theory in Eq. (2) admits vortex, or string solutions, in the broken (superconducting) phase. In other words, the classical equations of motion have solutions in which the symmetry is restored at the core,
but is broken far away from the core [24]. The existence of a vortex inside a loop \( C \) can be identified by computing the line integral

\[
n_C = \frac{1}{2\pi} \oint_C dx \cdot \nabla \gamma(x),
\]

where the winding number \( n_C \) is an integer, and a non-zero \( n_C \) signals a vortex inside the loop. It can be seen from Eq. (3) that \( n_C \) is gauge-invariant for single-valued gauge transformations \( \theta(x) \). Thus vortices are physical objects, which can be generated in a phase transition or by applying non-trivial boundary conditions, and they can be observed in the superconducting phase. However, in the equilibrium state they are also created and destroyed by thermal fluctuations, and \( \langle |n_C| \rangle \) gives their density both in the broken and in the symmetric phase. Moreover, it is expected that the behavior of vortices is qualitatively different in the two phases. As the dynamics is non-perturbative, it is then clear that it is very important to be able to study vortices with lattice simulations.

It should be noted here that one often assumes that in the regime of large \( x \), the gauge fields are not essential and one can approximate the theory in Eq. (2) by the 3d XY-model with a global U(1) symmetry, or its dual version (see, e.g., [13,26,27]) whose fundamental objects are the vortices. While these approximations simplify the problem significantly, their validity is uncertain for finite \( x \). Thus we consider it essential to approach the problem directly with the original theory in Eq. (2).

To allow for lattice simulations, the theory in Eq. (2) has to be discretized. As usual, we introduce the link field \( U_i(x) = \exp\{i\alpha_i A_i(x)\} \equiv \exp\{i\alpha_i(x)\} \). Rescaling the continuum scalar field to a dimensionless lattice field by \( \phi \to \beta_\mu \phi, \phi' \to 2a, \) the lattice action becomes

\[
S = \beta_\mu \sum_{x,i<j} \frac{1}{2} \tilde{F}_{ij}^2(x) - \beta_H \sum_{x,i} \text{Re} \phi^* (x) U_i(x) \phi(x + \hat{i}) + \sum_x \phi^*(x) \phi(x) + \beta_R \sum_x \left[ \phi^*(x) \phi(x) - 1 \right]^2,
\]

where \( \tilde{F}_{ij}(x) = \alpha_i(x) + \alpha_i(x + \hat{j}) - \alpha_i(x + \hat{j}) - \alpha_i(x) \). We use here the non-compact formulation for the gauge fields. The gauge transformation properties in Eq. (3) go over into

\[
\alpha_i(x) \to \alpha_i(x) + \theta(x) - \theta(x + \hat{i}).
\]

Discretization can be viewed as a different regularization scheme, and in order to describe the same continuum physics as in Eq. (2), one has to make a 2-loop computation relating the counterterms [28]. As a result, the lattice couplings \( \beta_\mu, \beta_H, \beta_R \) are determined from

\[
\beta_\mu = \frac{1}{e_3^2 a}, \quad \beta_R = \frac{x \beta_H^2}{4 \beta_\mu}, \quad 2 \beta_H \left( \frac{1}{\beta_H} - 3 - \frac{x \beta_H}{2 \beta_\mu} \right) = y - \frac{3.175115(1 + 2 x) \beta_G}{2\pi} - \frac{(-4 + 8 x - 8 x^2) \log \beta_G + 0.09}{16 \pi^2} + \frac{1.1 - 4.6 x}{16 \pi^2}.
\]

Thus for a given continuum theory depending on one scale \( e_3^2 \) and the two dimensionless parameters \( y, x \), the use of a lattice introduces a regulator scale \( a \), and Eqs. (6)–(9) specify, up to terms of order \( e_3^2 a \), the corresponding lattice action. Note that the simulations in [29] correspond to \( \beta_R \to \infty \), whereas according to Eq. (8), \( \beta_R \to 0 \) in the continuum limit for any finite \( x \) (\( \beta_H \to 1/3 \)).

3. Gauge-invariant vortices

Consider now vortices. The naive discretization of Eq. (5) gives the standard algorithm used in scalar theories without gauge fields. For each loop \( C \) one would define the winding number \( \tilde{n}_C \) of the phase \( \gamma \) of the scalar field. However, any \( \gamma(x) \) can be changed arbitrarily with a gauge transformation, see Eq. (7). Thus the \( \gamma(x) \)'s are essentially random numbers, and \( \tilde{n}_C \) does not contain any real informa-
tion about the dynamics of the system. Indeed, the result for a loop around a single plaquette, $\tilde{N}_{1 \times 1} = \langle |\tilde{n}_{1 \times 1}(x)| \rangle$, would equal $1/3$ in the case of completely uncorrelated field values, and this is what we measure from lattice simulations for $\tilde{N}_{1 \times 1}$, irrespective of the parameters of the theory (to be more precise, we always get $\tilde{N}_{1 \times 1} = 0.32(1) \ldots 0.33(1)$). Thus the quantity $\tilde{n}_c$ has to be rejected.

One solution sometimes used in the literature would be to fix the gauge, which makes $\tilde{n}_c$ non-trivial. However, its value depends crucially on the gauge chosen, and it is even possible to choose a gauge in which $\tilde{n}_c$ always vanishes. Therefore, we believe that it is important to use an explicitly gauge-invariant definition, in order to be able to interpret the results correctly.

Fortunately, the problem with $\tilde{n}_c$ is not one of principle, and a satisfactory definition can be given. For each positively directed link $l = (x, x + \hat{i})$ let us define

$$Y_{(x, x + \hat{i})} = [\alpha_i(x) + \gamma(x + \hat{i}) - \gamma(x)]_\sigma - \alpha_i(x).$$

(10)

For links with negative direction the sign of $Y_i$ is changed: $Y_{(x, x - \hat{i})} = -Y_{(x - \hat{i}, x)}$. Then, for each closed loop $C$, we can define

$$Y_c = \sum_{i \in C} Y_i \equiv 2\pi n_c.$$  

(11)

This definition has four main properties:

(a) For any field configuration and any loop $C$, $n_c \in \mathbb{Z}$.

(b) Directly from Eqs. (7), one can see that the part of $Y_i$ in the square brackets is gauge-invariant. The term $-\alpha_i(x)$ is not, but when summed over a closed loop into $Y_c$, the gauge dependence cancels. Hence $Y_c$ is gauge-invariant.

(c) Since $Y_c$ is gauge-invariant, one can always tune the gauge used in the evaluation of the $Y_i$’s such that the fields appearing are perturbatively small. But then, in the continuum limit, $\alpha_i = e_i/A$, goes to zero and one gets the correct continuum limit containing only the phase of $\phi$.

(d) The quantity $Y_c$ is additive: if there is a loop $C$ consisting of the loops $A, B$, then $Y_c = Y_A + Y_B$. This is what one would require for counting the number of vortices going through loops of different sizes. This also implies that vortex lines cannot end, and therefore they form closed vortex loops. Note that we use a non-compact gauge field so that there are no monopoles.

Based on these properties, Eq. (11) provides a valid formulation for counting vortices in the locally symmetric U(1) theory [30]. This definition of the winding number coincides with that given in Ref. [29] for the case $\beta_R = \infty$.

4. Simulations and results

In order to see how the gauge-invariant definition performs in practice, we have made lattice Monte Carlo simulations in the GL model. We choose $C = n \times n \equiv a$ loop around a plaquette of size $n \times n$ in Eq. (11), and measure

$$N_{n \times n} \equiv \langle |n_{n \times n}| \rangle.$$  

(12)

The quantity $N_{n \times n}$ measures the average net number of vortices through a loop of size $n \times n$, irrespective of the net direction. Keeping track of the direction would give zero: for symmetry reasons, $\langle |n_{n \times n}| \rangle = 0$. In practice, we average $\langle |n_{n \times n}| \rangle$ over all lattice sites and directions, to improve on the statistics.

Simulations are made at two values of $x$: $x = 0.0463$ corresponds to a strongly type I superconductor, $x = 2$ to a strongly type II superconductor. For each $x$, values of $y$ are chosen on both sides of the transition (see Fig. 1). For each such continuum parameter point, several lattice spacings are chosen: $\beta_G = 1/(e_0^2 a) = 1, 2, 3, 4, 6, 8, 12$. For $\beta_G = 4$ ($x = 0.0463$) and $\beta_G = 1.2$ ($x = 2$), several volumes are chosen, in order to test that the finite volume effects are small. The volumes thus arrived at ($24^2 \times 48$ for $\beta_G = 4$ both at $x = 0.0463, x = 2$) are then scaled with $\beta_G$ such that the physical volume (in units of $1/e_0^2$) remains constant. The orders of magnitude for $\beta_G$ and the volume come from the requirement that the physical correlation lengths, of order $(0.5 \ldots 2)/e_0^2$ at $x = 0.0463$ and $(1.5 \ldots 3)/e_0^2$ at $x = 2$ [17], are much longer than the lattice spacing but much shorter than the extent of the whole lattice (with the exception of the photon in the symmetric phase).

The results for $N_{1 \times 1}$ as a function of $y$ are shown in Fig. 2. The two values of $x$ have a different
corresponds precisely to the continuum limit. The point of a second order phase transition, which is close to the "universal" value seen, e.g., in 20 whereas in the regime of large $x$, $N_{1 \times 1}$ behaves continuously. Some of the points at $x = 0.0463$ correspond to metastable phases.

critical point $y_c$, see Fig. 1. At small $x$ the vortex density is very small in the broken phase but jumps discontinuously to a large value at the transition. However, the symmetric-phase value $= 0.2$ is much smaller than the trivial value $1/3$. At large $x$ the behavior is completely different. The total vortex density is large also rather deep in the broken phase and there is no discontinuity at the transition.

Let us then discuss the approach to the continuum limit: $a \to 0$ and hence $\beta_G \to \infty$. The continuum extrapolation of $N_{1 \times 1}$ is shown in Fig. 3. It is seen that even though there is a lot of structure at a finite $\beta_G$, all the structure disappears when $\beta_G \to \infty$ and one gets a result which is independent of the parameters of the continuum theory [31]. In the continuum limit, the loop of size $1 \times 1$ (as well as a loop of any finite size $n \times n$ in lattice units) shrinks to a point, and it is clear that the result is merely an artifact of the lattice regularization, and sensitive only to UV-effects. Note that the continuum value $N_{1 \times 1} \sim 0.2$ is close to the "universal" value seen, e.g., in [20] (there it was obtained at a fixed lattice spacing but at the point of a second order phase transition, which corresponds precisely to the continuum limit).

An important point to be noticed from Fig. 3 is that in the broken phase ($x = 0.0463, y = 0.08$), the asymptotic $\beta_G$ regime is obtained quite late, $\beta_G \geq 8$. This is somewhat surprising since the smallest physical correlation length at this point is $1/m_H \sim 0.4/\epsilon_f^2$ [17], corresponding to $\sim 2.4a$ already at $\beta_G = 6$.

Thus one would expect to be approaching the continuum limit earlier. The Higgs correlation length is much larger, $1/m_H \sim 2/\epsilon_f^2 \sim 12a$ at $\beta_G = 6$.

For larger loops $n \times n$, the qualitative behaviour is similar to that for $N_{1 \times 1}$, although the numerical values are different. For large $\beta_G$, including only terms linear in $a$, we expect the values of $N_{n \times n}$ to behave as

$$N_{n \times n} = f(n) + \left[ d(x, y) + e(x, y) n \right]/\beta_G + \sigma(1/\beta_G^q).$$

(13)

In principle, there could be a $\ln \beta_G$-term in $d(x, y), e(x, y)$. For fixed $n$, the continuum value is $f(n)$, but it does not reflect the infrared dynamics of the theory. Fitting the data, the functional form of $f(n)$ is found to be consistent with $a + b/n + \epsilon \ln n$ for large $n$. We cannot conclusively determine whether the coefficient $\epsilon$ is non-vanishing or not. Assuming $\epsilon = 0$, we get $a \sim 0.33, b \sim -0.13$, but if $\epsilon$ is allowed to be non-zero, the absolute values of $a$ and $b$ are somewhat smaller. The numerical determi-
nation of $\bar{c}$ is difficult, since it requires large values of $n$, for which very large lattices are needed to remove the finite size effects. In any case, the real physics lies in the coefficient of $1/\beta_0$ (see below), for which a fit of the form in Eq. (13) works very well (the confidence level is $CL = 10-90\%$, depending on the parameter values $x, y$).

Let us then discuss loops which are of a fixed size in physical units: $(c/e_3^2) \times (c/e_3^2)$, where $c$ is a constant. In lattice units this loop is $c \beta_0 \times c \beta_0$, i.e., the size is varied as a function of $\beta_0$. The continuum extrapolation for the loop $N_{\beta_0 \times \beta_0}$ ($c = 1$) is shown in Fig. 4 at the point $x = 0.0463, y = 0.14$. According to Eq. (13), we expect that in the continuum limit,$$
abla_{(c \beta_0 \times (c \beta_0)} \approx \lim_{\beta_0 \to 0} f(c \beta_0) + ce(x, y) + \sigma(e^2).

(14)

The first term is unphysical and corresponds to the regularization effects in Eq. (15) below. (If there is a term $\ln \beta_0$ in $f(x, y)$, then the regularization sensitive part can also depend on $x, y$, but this dependence is analytic and does not affect any phase transitions.) Thus the absolute value of $N_{(c \beta_0 \times (c \beta_0)}$ is not physical, only its changes are (see Fig. 5). The physical effects, i.e. $ce(x, y) + \sigma(e^2)$, come from the coefficient of $1/\beta_0$ (and $\sigma(1/\beta_0^2)$) in Eq. (13). To understand better the behaviour in Eq. (14), let us discuss observables simpler than $n_c$.

Consider a typical composite operator, such as $\langle \phi \cdot \phi \rangle$. For $\langle \phi \cdot \phi \rangle$, one can make a perturbative computation to find out what happens in the continuum limit. It turns out there is a linear (1-loop) and logarithmic (2-loop) divergence. The finite MS-scheme continuum result $\langle \phi \cdot \phi(e_3^2) \rangle_{\text{cont}}$ is [28]

$$\langle \phi \cdot \phi(e_3^2) \rangle_{\text{cont}} = \frac{e_3^2}{2} \beta_0 \beta_0 \langle \phi \cdot \phi \rangle_{\text{latt}} - \frac{3.1759115 \beta_0}{4 \pi} - \frac{1}{8 \pi^2} \log(6 \beta_0) + 0.668,$

(15)

where “latt” refers to the normalisation of the field in the lattice action (6). The second term on the RHS is the linear, and the third the logarithmic divergence. The value of $\langle \phi \cdot \phi \rangle_{\text{latt}}$ measured on the lattice is thus not in itself a physical quantity; only its changes (such as the discontinuity across a first order transition) are, since in them the divergent parts cancel.
We now expect that a similar thing happens for \( n_c \). The difference is that \( n_c \) is a more complicated and non-local quantity. The regularization sensitive part is not easily computable in perturbation theory, since even the integral

\[
N^\text{free} = \int dD \phi \left| n_c \right| \exp \left( -\int d^3 x |\partial_x \phi|^2 \right)
\]

is not Gaussian, due to the non-polynomial expression of \( |n_c| \). A numerical evaluation of \( N^\text{free} \) (a mass term has to be included on the lattice to kill a zero mode) gives a result close to the fitted values of \( f(n) \) in Eq. (13) (and favours \( \tilde{c} = 0 \)).

To demonstrate that the changes in \( N_{\beta_G} \) are physical, the distributions of \( N_{\beta_G} \) at \( x = 0.0463, \beta_G = 2,4,8 \) around the first order phase transition are shown in Fig. 5. It is seen that the two-peak structure (in particular, the distance between the peaks) indeed remains the same within statistical errors when \( \beta_G \) is varied, when \( \beta_G \) is large enough \( (\beta_G = 4,8) \). Thus the two-peak structure is physical and has a finite continuum limit, while the location of the structure on the \( |n| \)-axis is unphysical and dominated by UV-effects: both peaks move to the right when \( \beta_G \) increases (the location of the \( y = 0.14 \) peak is shown in Fig. 4). Note that \( \beta_G = 2 \) is not yet in the scaling regime, and thus the distance between the peaks is different from that at \( \beta_G = 4,8 \).

For \( x = 2 \), there is only a single peak which moves continuously to larger values as \( y \) is increased, see Fig. 2.

5. Conclusions

In conclusion, we have given a gauge-invariant definition for a vortex passing through a loop on a lattice, measuring the corresponding total vortex density, and discussed its extrapolation to the continuum limit. We have pointed out that to approach a meaningful continuum limit, one must keep the size of the loops fixed in physical units. We have found that the total vortex density behaves as a pseudo order parameter, analogously to \( \langle \phi^2 \rangle \): the absolute value is always non-zero and is dominated by regularization effects near the continuum limit. Thus only the changes of the total vortex density with respect to the continuum parameters are physically meaningful.

In the type I regime, the total vortex density displays a discontinuity, whereas in the type II regime, it behaves smoothly as the phase transition is crossed (Fig. 2).

The system possesses also (non-local) observables which behave analogously to true order parameters. One such is the photon mass, which vanishes exactly in the symmetric phase. It has been suggested that another such quantity might be the density of long vortices passing through the whole lattice \([25,20]\). In contrast to the photon mass, this quantity should vanish in the broken phase and remain non-zero in the symmetric phase. The measurement of the density of long vortices is in progress \([19]\).

Finally, it would be interesting to study the spatial distribution of vortices. This can be done by measuring correlators of ‘‘vortex density operators’’ \( n_c \), separated by a distance \( r \). One can define several correlators, depending on the relative orientations of the loops used in the \( n_c \)-s. These quantities would give realistic initial conditions for simulations of the time evolution of vortex networks.

Acknowledgements

A.R. would like to thank G. Blatter and T.W.B. Kibble for discussions. We are grateful to G.D. Moore and K. Rummukainen for very useful comments. The simulations were carried out with a Cray C94 at the Finnish Center for Scientific Computing. This work was partly supported by the TMR network Finite Temperature Phase Transitions in Particle Physics, EU contract no. FMRX-CT97-0122.

References

Exact chiral symmetry on the lattice and the Ginsparg-Wilson relation

Martin Lüscher

Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, D-22603 Hamburg, Germany

Received 20 February 1998
Editor: P.V. Landshoff

Abstract

It is shown that the Ginsparg-Wilson relation implies an exact symmetry of the fermion action, which may be regarded as a lattice form of an infinitesimal chiral rotation. Using this result it is straightforward to construct lattice Yukawa models with unbroken flavour and chiral symmetries and no doubling of the fermion spectrum. A contradiction with the Nielsen-Ninomiya theorem is avoided, because the chiral symmetry is realized in a different way than has been assumed when proving the theorem. © 1998 Elsevier Science B.V. All rights reserved.

1. A well-known problem with fermions on the lattice is that one usually ends up with breaking chiral symmetry or having more particles in the continuum limit than intended. The celebrated Nielsen-Ninomiya theorem [1] states that this is in fact unavoidable if a few plausible assumptions are made. The construction of chiral field theories on the lattice thus appears to be difficult and maybe even impossible in some cases.

Recently some intriguing results have been published by Neuberger [2,3] and by Hasenfratz, Laliena and Niedermayer [4], which suggest that chiral symmetry may be preserved in lattice QCD, at least to some extent, if the lattice Dirac operator is of a particular form. The proposed expressions for the Dirac operator have been derived in completely different ways and tend to be very complicated, but all of them satisfy a simple identity, originally due to Ginsparg and Wilson [5], which protects the quark masses from additive renormalizations [2,6] and which plays a key rôle in the proof of the lattice index theorem of Ref. [4].

In this letter it will be shown that the lattice fermion action in fact has an exact symmetry if the Ginsparg-Wilson identity holds. The usefulness of this observation is illustrated by constructing a class of chiral Yukawa models on the lattice with unbroken chiral and flavour symmetries (and no doublers). Since the chiral transformations in these theories are not of the naively expected form, a contradiction with the Nielsen-Ninomiya theorem is avoided without having to compromise in any other ways. In particular, the flavour-singlet chiral symmetry has
the expected anomaly if gauge interactions are included.

2. A particularly simple form of the Nielsen-Ninomiya theorem holds for free Dirac fermions on a euclidean lattice and it is helpful for the discussion that follows to briefly recall this. So let us consider the free field action

\[ S = a^4 \sum_x \bar{\psi} D\psi, \]  

(2.1)

where \( a \) denotes the lattice spacing and \( D \) the lattice Dirac operator. As usual we assume \( D \) to be invariant under translations so that

\[ D e^{i\psi u} = \bar{D}(p)e^{i\psi u} \]  

(2.2)

for all constant Dirac spinors \( u \) and some complex \( 4 \times 4 \) matrix \( \bar{D}(p) \). The theorem now states that the following properties cannot hold simultaneously for \( D \):

1. \( \bar{D}(p) \) is an analytic periodic function of the momenta \( p_\mu \) with period \( 2\pi/a \).
2. For momenta far below the cutoff \( \pi/a \), we have \( \bar{D}(p) = i\gamma_\mu p_\mu \) up to terms of order \( ap^2 \).
3. \( \bar{D}(p) \) is invertible at all non-zero momenta (mod \( 2\pi/a \)).
4. \( D \) anti-commutes with \( \gamma_5 \).

Property (1) is necessary if we want \( D \) to be an essentially local operator, (2) and (3) ensure that the correct continuum limit is obtained and (4) guarantees that the fermion action is invariant under continuous chiral transformations 2.

3. To escape the theorem, Ginsparg and Wilson [5] suggested many years ago to replace property (4) through the relation

\[ \gamma_5 D + D\gamma_5 = \delta \gamma_5 D. \]  

(3.1)

A simple consequence of this equation is that the fermion propagator anti-commutes with \( \gamma_5 \) at non-zero distances and chiral symmetry is thus partly preserved.

Examples of free lattice Dirac operators satisfying the Ginsparg-Wilson relation can be found rather easily. A particularly simple solution is given by [2]

\[ D = \frac{1}{a}(1 - A(A')^{-1/2}), \quad A = 1 - aD_w, \]  

(3.2)

where \( D_w \) denotes the standard Wilson-Dirac operator,

\[ D_w = \frac{1}{2}\left(\gamma_\mu \tilde{\gamma}_\mu + \tilde{\gamma}_\mu \gamma_\mu\right) - a\gamma_u \tilde{\gamma}_u \]  

(3.3)

and \( \tilde{\gamma}_\mu \) and \( \tilde{\gamma}_\mu \) are the nearest-neighbour forward and backward difference operators. Because of the square root in Eq. (3.2), one might assume that \( D \) is a non-local operator, but this is actually not the case. Using the abbreviations \( p_\mu = (1/a)\sin(ap_\mu) \) and \( \tilde{p}_\mu = (2/a)\sin(ap_\mu/2) \), we have

\[ a\bar{D}(p) = 1 - \left(1 - \frac{i}{2}a^2 \tilde{p}^2 - ia\gamma_\mu \tilde{p}_\mu\right) \times \left(1 + \frac{1}{2}a^4 \sum_{\mu < \nu} \tilde{p}_\mu^2 \tilde{p}_\nu^2 \right)^{-1/2}, \]  

(3.4)

and it is immediately clear from this formula that the conditions (1), (2) and (3) listed above are fulfilled. In particular, from the analyticity of \( \bar{D}(p) \) one infers that its Fourier transform falls off exponentially at large distances with a rate proportional to \( 1/a \). For free fermions Eq. (3.2) thus provides a completely satisfactory solution of the Ginsparg-Wilson relation.

4. We now show that Eq. (3.1) implies a continuous symmetry of the fermion action, which may be interpreted as a lattice form of chiral symmetry. No particular assumptions need to be made here, i.e. the discussion applies to any Dirac operator satisfying the Ginsparg-Wilson relation, including the gauge covariant operators of Refs. [2–4].

The infinitesimal variation of the fields associated with the new symmetry is

\[ \delta \psi = \gamma_5 (1 + \frac{i}{2}aD) \psi, \quad \delta \bar{\psi} = \bar{\psi}(1 + \frac{i}{2}aD) \gamma_5. \]  

(4.1)

where \( D \) is considered to be a matrix which may be multiplied from the right with \( \psi \) or from the left with \( \bar{\psi} \). Flavour non-singlet chiral transformations may be defined similarly by including a group generator in Eq. (4.1). In both cases it is trivial to check that the fermion action Eq. (2.1) is invariant if the Ginsparg-Wilson identity holds.

The flavour-singlet chiral symmetry is anomalous in the presence of gauge fields and it now seems that
we have got too much symmetry on the lattice. The paradox is resolved by noting that the fermion integration measure is in general not invariant under the transformation (4.1).

To work this out let us consider the theory in a finite space-time volume with suitable boundary conditions so that the Ginsparg-Wilson identity is preserved. We are then interested in the symmetry properties of the (unnormalized) expectation values

$$\langle \sigma \rangle_F = \int \prod_x d\psi (x) d\overline{\psi}(x) \sigma e^{-S_F}$$

(4.2)
of arbitrary products $\sigma$ of the fermion fields. By substituting

$$\psi \to \psi + \epsilon \delta \psi, \quad \overline{\psi} \to \overline{\psi} + \epsilon \delta \overline{\psi},$$

(4.3)
and expanding to first order in $\epsilon$ one obtains

$$\langle \delta \sigma \rangle_F = - a \text{tr} \{ \gamma_5 D \langle \sigma \rangle_F \},$$

(4.4)
where the trace is to be taken over the space of all fermion fields. Evidently in the case of free fermions, with $D$ as given above, the trace vanishes and the symmetry is exact. The same is also true if we consider flavour non-singlet chiral rotations, because the group generator which has to be included in the transformation law and which then appears on the right-hand side of Eq. (4.4) is traceless.

The anomaly, $- a \text{tr} \{ \gamma_5 D \}$, has previously been calculated in Ref. [4] and we now give a second derivation which is applicable also in those cases where the Dirac operator does not have any particular hermiticity properties. Let $z$ be any complex number not contained in the spectrum of $D$. A little algebra, using the Ginsparg-Wilson identity, yields

$$a(z - D) \gamma_5 (z - D) = z(2 - az) \gamma_5$$

$$- (1 - az) \{ (z - D) \gamma_5 + \gamma_5 (z - D) \},$$

(4.5)
and after multiplying this equation from the right with $(z - D)^{-1}$ and taking the trace one ends up with

$$- a \text{tr} \{ \gamma_5 D \} = z(2 - az) \text{tr} \{ \gamma_5 (z - D)^{-1} \}.$$ (4.6)

We now divide through the factor $z(2 - az)$ and integrate over a small circle centred at the origin that does not encircle any spectral value of $D$ other than 0. In particular,

$$P_0 = \oint \frac{dz}{2\pi i} (z - D)^{-1}$$

(4.7)
projects on the subspace of zero modes of $D$ and the result

$$- a \text{tr} \{ \gamma_5 D \} = 2 \text{tr} \{ \gamma_1 P_0 \} = 2N_f \times \text{index}(D)$$

(4.8)
is thus obtained, where $N_f$ denotes the number of fermion flavours. Taken together Eqs. (4.4) and (4.8) show that the Ward identities associated with the global flavour-singlet chiral transformations on the lattice have the correct anomaly. In view of the exact index theorem of Hasenfratz et al. [4] and the earlier work of Ginsparg and Wilson [5] this comes hardly as a surprise, but it is striking that the anomaly can be calculated with so little effort.

5. It is now relatively easy to couple fermions to scalar fields in such a way that the flavour and chiral symmetries are preserved on the lattice. Our starting point is the free fermion action

$$S_F = a^4 \sum_x \{ \overline{\psi} D \psi - (2/a) \overline{\chi} \chi \},$$

(5.1)
where $D$ is assumed to be a decent solution of the Ginsparg-Wilson relation such as the one discussed in Section 3. The auxiliary fields $\chi$ and $\overline{\chi}$ will later be used to construct chirally invariant interaction terms. For the time being we only note that they do not propagate and the physical content of the theory is hence unchanged.

As before one can show that the modified transformation

$$\delta \psi = \gamma_5 (1 - \frac{i}{2} a D) \psi + \gamma_5 \chi, \quad \delta \chi = \gamma_5 \frac{i}{2} a D \psi,$$

$$\delta \overline{\psi} = \overline{\psi} (1 - \frac{i}{2} a D) \gamma_5 + \overline{\chi} \gamma_5, \quad \delta \overline{\chi} = \overline{\chi} \frac{i}{2} a D \gamma_5,$$

(5.2)
leaves the action and the fermion integration measure invariant (gauge interactions are excluded in this section). It follows from these equations that

$$\delta (\psi + \chi) = \gamma_5 (\psi + \chi), \quad \delta (\overline{\psi} + \overline{\chi}) = (\overline{\psi} + \overline{\chi}) \gamma_5,$$

(5.3)
and the propagator of the sum $\psi + \chi$ is hence chirally invariant in the ordinary sense. This is, incidentally, perfectly consistent with the Nielsen-Ninomiya theorem, because the Fourier transform of the propagator vanishes at some momenta and its inverse is hence singular, thus violating property (1) (cf. Section 2).
Suppose now that \( \phi \) is a complex scalar field on the lattice with the usual self-interactions. A chirally invariant Yukawa interaction term is then given by

\[
S \propto g_0 \sum_x (\bar{\psi} + \bar{\chi}) \left\{ \frac{i}{2} (1 - \gamma_5) \phi + \frac{i}{2} (1 + \gamma_5) \phi^* \right\} \times (\psi + \chi),
\]

with \( g_0 \) being the bare coupling constant. More complicated interactions with flavour symmetries and various multiplets of fermions can be constructed similarly. A few remarks should be added at this point to make it clear that the lattice theories defined in this way are completely sane. For simplicity attention is restricted to the phase where chiral symmetry is not spontaneously broken.

1. In perturbation theory the one-particle irreducible diagrams are chirally invariant in the ordinary sense, because the internal fermion lines represent the propagation of \( \psi + \chi \) and the vertices are manifestly invariant. Non-symmetric counterterms are hence not needed to renormalize the theory.

2. At sufficiently weak coupling the spectrum of fermions is exactly as expected, i.e. there are no doublers. To see this first note that there are none when the interactions are switched off. Now since the one-particle irreducible self-energy diagrams anti-commute with \( \gamma_5 \), they must be proportional to \( \gamma_5 P_\mu \) at small momenta. In particular, the perturbative corrections to the fermion propagator are of the form \( D^{-1} B \) where \( B \) is bounded and it is hence impossible that new poles arise at small couplings.

3. The auxiliary fields \( \chi \) and \( \bar{\chi} \) couple only locally and do not carry any independent physical information. Essentially these fields play the rôle of Lagrange multipliers which may integrated out if so desired although the expressions that one obtains are not particularly illuminating.

6. The important qualitative message of this paper is that the Nielsen-Ninomiya theorem can be bypassed if we do not insist that the chiral transformations assume their canonical form on the lattice. The construction of chirally invariant lattice theories remains non-trivial, however, because the transformation laws depend on the interaction in general. This is in fact required in gauge theories as otherwise one would end up with a non-anomalous flavour-singlet chiral symmetry. An interesting observation in this connection is that the chiral rotations \( (5.2) \) become interaction dependent when the auxiliary fields are eliminated.

At this point many interesting questions have not even been touched and are left for future research. In particular, the precise conditions under which chiral symmetries can exist on the lattice remain to be uncovered and an attempt should be made to derive an identity, similar to the Ginsparg-Wilson relation, which allows one to construct exactly supersymmetric lattice theories.

Acknowledgements

Helpful correspondence with Karl Jansen, Peter Hasenfratz, Ferenc Niedermayer and Herbert Neuberger is gratefully acknowledged. I am also indebted to Istvan Montvay for sharing his views on chiral Yukawa models.

References

Various properties of compact QED and confining strings

D.V. Antonov 1, 2

Institut für Physik, Humboldt-Universität zu Berlin, Invalidenstrasse 110, D-10115 Berlin, Germany

Received 26 February 1998
Editor: P.V. Landshoff

Abstract

The effects bringing about by the finiteness of the photon mass due to the Debye screening in the monopole gas in three-dimensional compact QED are studied. In this respect, a representation of the partition function of this theory as an integral over monopole densities is derived. Dual formulation of the Wilson loop yields a new theory of confining strings, which in the low-energy limit almost coincides with the one corresponding to the case when the photon is considered to be massless, whereas in the high-energy limit these two theories are quite different from each other. The confining string mass operator in the low-energy limit is also found, and its dependence on the volume of observation is studied. © 1998 Elsevier Science B.V. All rights reserved.

PACS: 11.10.Kk; 11.10.Lm; 11.15.-q; 11.25.Sq
Keywords: Quantum electrodynamics; Sine-Gordon model; String model; Effective potential; Saddle-point approximation

1. Introduction

Theory of magnetic monopoles in compact QED both on a lattice and in the continuum limit has a long history [1–3]. In Ref. [2], it has been demonstrated that at large distances the partition function of 3D compact QED is nothing else but the one of a gas of monopoles with a Coulomb interaction, which in notations of Ref. [2] reads

\[ Z = \sum_{N=1}^{+\infty} \sum_{q_a = \pm 1} \frac{\xi^N}{N!} \prod_{i=1}^N \int dz_i \exp \left[ -\frac{\pi}{2e^2} \sum_{a \neq b} \frac{q_a q_b}{|z_a - z_b|^2} \right] \] (1)

1 E-mail address: antonov@pha2.physik.hu-berlin.de.
2 On leave of absence from the Institute of Theoretical and Experimental Physics, B.Cheremushkinskaya 25, 117 218, Moscow, Russia.

0370-2693/98/$19.00 © 1998 Elsevier Science B.V. All rights reserved.
PII: S0370-2693(98)00419-5
Due to Ref. [2], Eq. (1) could be equivalently rewritten as a partition function of a 3D sine-Gordon model of a scalar field, which effectively substitutes an infinite number of monopoles,

\[ Z = \int \mathcal{D} \chi \exp \left( - \int dx \left( \frac{1}{2} (\partial_x \chi)^2 - 2 \zeta \cos \left( \frac{2\pi}{e} \chi \right) \right) \right). \tag{2} \]

Universal Confining String Theory (UCST), proposed in Ref. [3], is a dual formulation of the Wilson loop in the model (2). This formulation enables one to rewrite the interaction of the field \( \chi \) with a solid angle, formed by the point of observation and the contour \( C \) of the Wilson loop, as an interaction of an antisymmetric tensor field, possessing a multivalued action, with the string world-sheet. In this respect, summation over surfaces bounded by \( C \), i.e. summation over string world-sheets, could be treated as a summation over the branches of this multivalued action.

However, as it has already been discussed in Ref. [2], due to the Debye screening in the monopole gas, the photon in the model under study acquires a mass equal to the mass of the field \( \chi \), following from the low-energy expansion of Eq. (2), \( m = \frac{\zeta}{2} \sqrt{2 \zeta} \). This means that this mass should be taken into account in the monopole interaction from the very beginning, i.e. the Coulomb interaction in Eq. (1) should be replaced by the Yukawa one with the mass \( m \), which leads to the following modification of Eq. (1)

\[ Z = \sum_{N=1}^{+\infty} \sum_{q=\pm 1} \frac{\zeta^N}{N!} \prod_{r=1}^{N} \int dz_r \exp \left[ -\frac{\pi}{2 e^2} \sum_{a \neq b} \frac{q_a q_b}{|z_a - z_b|} e^{-m |z_a - z_b|} \right]. \]

This expression could be again represented as a partition function of the sine-Gordon model of an auxiliary scalar field with the mass \( m \sqrt{2} \),

\[ Z = \int \mathcal{D} \varphi \exp \left( - \int dx \left( \frac{1}{2} (\partial_x \varphi)^2 + \frac{m^2}{2} \varphi^2 - 2 \zeta \cos \left( \frac{2\pi}{e} \varphi \right) \right) \right). \tag{3} \]

In the next section, we shall study the effects bringing about by the additional mass term in Eq. (3) both to the 3D compact QED and UCST. First, we shall cast Eq. (3) into the form of an integral over monopole densities, after which we shall study the influence of the mass term to the UCST and treat the obtained string theory both in the low- and high-energy limits. The latter one differs drastically from such a limit corresponding to the case when the photon is considered to be massless.

In Section 3, we shall find the mass operator of a string in the low-energy limit of the UCST, corresponding to the 4D compact QED.

We shall end up with a short Conclusion.

2. Accounting for the finiteness of the photon mass in 3D Compact QED and UCST

Let us start with a representation of Eq. (3) as an integral over monopole densities. To this end, one should introduce into Eq. (3) a unity of the form

\[ \int \mathcal{D} \rho \delta \left( \rho(x) - \sum_a q_a \delta(x - z_a) \right) = \int \mathcal{D} \rho \mathcal{D} \mu \exp \left( i \int dx \mu \rho - \sum_a q_a \mu(z_a) \right), \]

which, after the change of variables, \( \mu = \frac{\zeta}{e} \varphi - \mu \), and integrations over the fields \( \varphi \) and \( \psi \), yields the following representation for the partition function

\[ Z = \int \mathcal{D} \rho \exp \left( -\frac{\pi}{2 e^2} \int dx dy \rho(x) \frac{e^{-m |x-y|}}{|x-y|} \rho(y) + V[\rho] \right). \]
where
\[ V[\rho] = \int dx \left( \rho \ln \left( 1 + \left( \frac{\rho}{2\xi} \right)^2 + \frac{\rho}{2\xi} \right) - 2\xi \ln \left( 1 + \left( \frac{\rho}{2\xi} \right)^2 \right) \right) \]

is a parabolic-type effective potential, whose asymptotic behaviours at \( \rho \ll \xi \) and \( \rho \gg \xi \) read
\[ V[\rho] \to \int dx \left( -2\xi + \frac{\rho^2}{4\xi} \right) \]

and
\[ V[\rho] \to \int dx \left( \rho \left( \ln \frac{\rho}{\xi} - 1 \right) \right) \]

respectively. Thus, one can represent 3D compact QED as a nonlocal theory of the monopole densities with the Yukawa interaction and a certain effective potential.

Let us now proceed to the UCST, i.e. to the dual formulation of the Wilson loop
\[ \langle W(C) \rangle = \left\langle \exp \left( i \sum_a q_a \eta(z_a) \right) \right\rangle = \int \mathcal{D} \varphi \mathcal{D} \varphi' \exp \left( - \int dx \left( \frac{1}{2} \left( \partial_\mu \varphi \right)^2 + \frac{m^2}{2} \varphi^2 - 2\xi \cos \left( \frac{2\pi}{e} \varphi + \eta \right) \right) \right) \]

where \( \eta(x) = \frac{1}{2} \int d\sigma(y) \frac{(x-y)}{(x-y)^2} \) stands for the solid angle formed by the point \( x \) and the contour \( C \). This could be done similarly to Ref. [3] by introduction of an auxiliary integration over an antisymmetric tensor field as follows:
\[ \langle W(C) \rangle = \int \mathcal{D} \phi \mathcal{D} B_{\mu\nu} \exp \left\{ - \int dx \left[ \frac{9}{e^2} B_{\mu\nu}^2 - \frac{3i}{e} \phi \epsilon_{\mu\nu\lambda} \partial_\lambda B_{\mu\nu} - 2\xi \cos \varphi + \xi (\varphi - \eta)^2 \right] + i \int d\sigma_{\mu\nu} B_{\mu\nu} \right\} \]

where \( \phi = \frac{2\pi}{e} \varphi + \eta \), \( d\sigma_{\mu\nu} = \epsilon_{\mu\nu\lambda} d\sigma_\lambda \), and further elimination of the field \( \phi \), which has no more kinetic term in the action. However now, contrary to Ref. [3], due to the presence of the mass term \( \xi \phi^2 \) of the field \( \phi \) in the Lagrangian, the saddle-point equation for this field has no more simple solutions and could be solved only in the low- or in the high-energy limits.

In the first case, i.e. when \( \phi \ll 1 \), the Wilson loop takes the following form:
\[ \langle W(C) \rangle = \exp \left\{ - \frac{\pi \xi}{12} \int d\sigma\mu\nu(x) \int d\sigma\mu\nu(y) \frac{1}{|x-y|} \right\} \]

\[ \times \int \mathcal{D} B_{\mu\nu} \exp \left\{ - \int dx \left( \frac{3}{16\pi^2} H_{\mu\nu\lambda}^2 + \frac{9}{e^2} B_{\mu\nu}^2 \right) + \frac{i}{2} \int d\sigma_{\mu\nu} B_{\mu\nu} \right\} \]

(4)

where \( H_{\mu\nu\lambda} = \partial_\lambda B_{\mu\nu} + \partial_\nu B_{\mu\lambda} + \partial_\mu B_{\nu\lambda} \) is the strength tensor of the field \( B_{\mu\nu} \). One can see that except for the \( B_{\mu\nu} \)-independent part of the action,
\[ \frac{\pi \xi}{12} \int d\sigma\mu\nu(x) \int d\sigma\mu\nu(y) \frac{1}{|x-y|} \]

(5)

Eq. (4) is just the low-energy limit of the UCST, obtained in Ref. [3] without accounting for the finiteness of the photon mass.

It is easy to calculate the string tension of the Nambu-Goto term and the inverse bare coupling constant of the rigidity term bringing about by Eq. (5). To this end, we shall make use of the results of Ref. [4] and introduce a
dimensionless cutoff $L = \frac{\alpha_s}{\Lambda^2} \sim \frac{1}{\varepsilon m}$, which is much larger than unity in the low-energy limit of the field $\varphi$ under consideration. Then the contributions of Eq. (5) to the quantities mentioned above read as follows:

$$\Delta \sigma = \frac{2}{3} \frac{\pi^2 k^3}{L^3} \sim \frac{\pi^2}{3} \frac{e^3}{L^3}$$

(6)

and

$$\Delta \frac{1}{\alpha_0} = - \frac{\pi^2}{144} L^3 \sim - \frac{\pi^2}{144} \frac{e^6}{\xi^2},$$

respectively.

Another case, when it is also possible to solve the saddle-point equation, is the case $\phi \gg 1$. This could be done by making use of the iterative procedure, which in the first order yields the following expression for the Wilson loop:

$$\langle W(C) \rangle = \int d^4 B_{\mu \nu} \exp \left[ - \int dx \left( \frac{1}{12} A^2 H_{\mu \lambda \nu} + \frac{1}{4} B_{\mu \nu}^2 - 2 \zeta \cos \left( \frac{i}{\Lambda^2} e_{\mu \nu} H_{\mu \nu} + \eta \right) \right) \right].$$

(7)

One can see that Eq. (7) is quite different from the expression for the UCST partition function corresponding to the case when the photon is considered to be massless, studied in Ref. [3].

3. Mass operator of the confining string in the low-energy limit

In the low-energy limit, the kinetic term of the field $B_{\mu \nu}$ in Eq. (4) vanishes, and we are left with the string described by the action (5), interacting with a constant antisymmetric tensor field with a certain Gaussian measure, which should be eventually averaged over. For simplicity, one can take into account only the Nambu-Goto term in the expansion of the action (5), whose string tension is given by Eq. (6).

The Nambu-Goto string interacting with a constant antisymmetric tensor field with a Gaussian measure appears also in the low-energy limit of the 4D UCST, studied in Ref. [5], where the finiteness of the photon mass has not been taken into account. In this case, it appears however not in the simplest approximation, when one treats the field $B_{\mu \nu}$ as a constant one. Namely, let us consider the low-energy expression for the partition function of the 4D UCST (i.e. the dual expression for the Wilson loop in the confining phase of 4D compact QED),

$$\langle W(C) \rangle = \int d^4 B_{\mu \nu} \exp \left[ - \int d^4 x \left( \frac{1}{12} A^2 H_{\mu \lambda \nu} + \frac{1}{4} B_{\mu \nu}^2 + i \int d \sigma_{\mu \nu} B_{\mu \nu} \right) \right].$$

where $A$ stands for the UV momentum cutoff, and split the total field $B_{\mu \nu}$ into an $x$-independent background part, $b_{\mu \nu}$, and a quantum fluctuation, $h_{\mu \nu}(x)$. Since due to the Hodge decomposition theorem [6], $h_{\mu \nu}(x)$ could be always represented in the form $h_{\mu \nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + \partial_\gamma C_{\mu \nu \gamma}$, where $A_\mu$ and $C_{\lambda \mu \nu}$ stand for some vector and an antisymmetric rank-3 tensor respectively, the term $\frac{1}{2\pi} \int d^4 x b_{\mu \nu} h_{\mu \nu}$ during this splitting vanishes by virtue of partial integration, and we arrive at the following expression for the Wilson loop:

$$\langle W(C) \rangle = \int_{-\infty}^{+\infty} \prod_{\mu, \nu=1, \mu \neq \nu}^4 db_{\mu \nu} \exp \left[ - \frac{V^{(4)} h_{\mu \nu}^2}{4 e^2} + i \int d \sigma_{\mu \nu} b_{\mu \nu} \right]$$

$$\times \exp \left[ - \int d \sigma_{\mu \nu} (x) \int d \sigma_{\mu \nu} (x') D_{\mu \nu} (x-x') \right].$$
where $V^{(4)}$ is the four-volume of observation, and $D_{a\beta,\lambda\gamma}(x-x')$ stands for the propagator of the massive Kalb-Ramond field $h_{\mu\nu}$. What is important for us here is not the explicit form of this propagator, but the fact that the leading term of the derivative expansion of the action $\int d\sigma_{a\beta}(x) d\sigma_{\gamma\lambda}(x') D_{a\beta,\lambda\gamma}(x-x')$ is the Nambu-Goto one with the string tension [5] $\sigma = \frac{1}{4\pi} K_0 \left( \frac{1}{4 \text{exp}(\frac{1}{\sqrt{\sigma}})} \right)$, where $K_0$ stands for the Macdonald function. Thus we have again arrived at a theory of the Nambu-Goto string interacting with the constant antisymmetric tensor field $b_{\mu\nu}$, which possesses the quadratic action $S[b_{\mu\nu}] = \frac{\sqrt{\text{det} b}}{4 \pi}$.

In what follows, for concreteness, we shall study this very theory, rather than the analogous one, described in the first paragraph of the present section, which follows from 3D compact QED, and find the mass operator of the confining string in this theory. To this end, we shall make use of the result of Ref. [7], where the mass operator of the Nambu-Goto string in the external constant electromagnetic field has been found. The difference of our case from the one studied in Ref. [7] is the necessity of performing the average over the field $b_{\mu\nu}$. In this way, we find the following expression for the operator of the square of mass of the confining string:

$$M^2 = 2\pi \sigma \left[ \sum_{n=1}^{+\infty} \frac{1}{n} \hat{a}_n^\dagger \hat{a}_n - \alpha(0) \right] \int_{-\infty}^{+\infty} \frac{2}{1 + \frac{2}{\sigma^2} \left( b_{12}^2 + b_{13}^2 + b_{14}^2 + b_{23}^2 + b_{24}^2 + b_{34}^2 \right)} \frac{\text{d}b_{\mu\nu}}{\text{d}b_{\mu\nu}} e^{-S[b_{\mu\nu}]},$$

(8)

where the eigenvalues of the operator $\sum_{n=1}^{+\infty} n \hat{a}_n^\dagger \hat{a}_n$ are equal to 1, 2, ..., and $\alpha(0) < 0$. In order to carry out the integral standing in the numerator on the R.H.S. of Eq. (8), one should split the infinite interval of integration over the absolute value of the six-dimensional vector into two pieces, from 0 to $\frac{\pi}{\sqrt{\sigma^2}}$ and from $\frac{\pi}{\sqrt{\sigma^2}}$ to $+\infty$, and neglect in these two intervals the terms $\frac{2}{\sigma^2} \left( b_{12}^2 + b_{13}^2 + b_{14}^2 + b_{23}^2 + b_{24}^2 + b_{34}^2 \right)$ and 1 in the integrand, respectively. Then Eq. (8) yields the following value of the operator of the square of mass of the confining string in the low-energy limit:

$$M^2 = 2\pi \sigma \left[ \sum_{n=1}^{+\infty} \frac{1}{n} \hat{a}_n^\dagger \hat{a}_n - \alpha(0) \right] \left( 1 - \frac{1}{2} \left( \frac{\sigma^2 V^{(4)}(4)}{8 e^2} \right)^2 + \frac{\sigma^2 V^{(4)}(4)}{8 e^2} + 1 \right) \exp \left( - \frac{\sigma^2 V^{(4)}(4)}{8 e^2} \right),$$

(9)

which could be checked to be always positive. In Eq. (9), the effect of the finiteness of the volume of observation has been taken into account. The value of $M^2$ in the limit $V^{(4)} \to +\infty$ is obvious.

4. Conclusion

In the present Letter, we have addressed two problems. In Section 2, we have studied the influence of the effects bringing about by the finiteness of the photon mass in the 3D compact QED to the representation of the partition function of this theory as an integral over monopole densities. As a result, we have obtained the nonlocal theory of these densities with the Yukawa interaction and a certain parabolic-type effective potential. Next, we have investigated the influence of the finiteness of the photon mass to the dual formulation of the Wilson loop, i.e. to the theory of confining strings. In the low-energy limit, the obtained string theory coincides (apart from the additional contribution (5) to the action) with the one corresponding to the case when the photon mass is not taken into account, whereas in the high-energy case, described by Eq. (7), these two theories are quite different from each other.
In Section 3, we have found the mass operator of the confining string corresponding to the low-energy limit of the 4D compact QED, where the photon has been considered to be massless. In the final result (9), the dependence on the four-volume of observation has been explicitly presented.

Acknowledgements

The author is indebted to Prof. H. Kleinert for bringing his attention to some Refs. of [1] and Dr. H. Dorn and V.I. Shevchenko for bringing his attention to Ref. [6]. He is also grateful to the theory group of the Quantum Field Theory Department of the Institut für Physik of the Humboldt-Universität of Berlin for kind hospitality. This work is supported by Graduiertenkolleg Elementarteilchenphysik, Russian Foundation for Basic Research, project 96-02-19184, DFG-RFFI, grant 436 RUS 113/309/0, and by the INTAS, grant No.94-2851.

References


Relation between QCD potentials in momentum and position space

M. Ježabek a,b, M. Peter c, Y. Sumino d,1

a Institute of Nuclear Physics, Kawiory 26a, PL-30055 Cracow, Poland
b Department of Field Theory and Particle Physics, University of Silesia, Uniwersytecka 4, PL-40007 Katowice, Poland
c Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, D-69120 Heidelberg, Germany
d Institut für Theoretische Teilchenphysik, Universität Karlsruhe, D-76128 Karlsruhe, Germany

Received 13 March 1998
Editor: P.V. Landshoff

Abstract

We derive a formula which relates the QCD potentials in momentum space and in position space in terms of the β function of the renormalization-group equation for the potential. This formula is used to study the theoretical uncertainties in the potential and in particular in its application to the determination of the pole mass m_p when we use perturbative expansions. We demonstrate the existence of these uncertainties for the Richardson potential explicitly and then discuss the limited theoretical accuracy in the perturbative QCD potential. We conclude that a theoretical uncertainty of m_p much below 100 MeV would not be achievable within perturbative QCD.

PACS: 12.38.-t; 12.38.Bw; 12.39.Pn

In this article we discuss a relation between the static QCD potential [1] in momentum space

\[ V(q) = -C_F \frac{4\pi \alpha_s(q)}{q^2}, \quad (1) \]

where for quarks \( C_F = 4/3 \), and the corresponding potential in position space 2

\[ \widetilde{V}(r) = \int \frac{d^3q}{(2\pi)^3} V(q) e^{iq \cdot r} = -C_F \frac{\pi \alpha_s(1/r)}{r}. \quad (2) \]

Eqs. (1) and (2) are understood as the definitions of the couplings \( \alpha_s(q) \) and \( \pi \alpha_s(1/r) \). In addition, they express the fact that these two functions are related because \( \widetilde{V}(r) \) is the Fourier transform of \( V(q) \). We can obtain a useful expression for this functional relation

\[ \pi \alpha_s(1/r) = F [\alpha_s(q)] \quad (3) \]

using the renormalization-group equation of the coupling \( \alpha_s(q) \), and we will study the consequences of this relation.

First we relate the coupling at a general scale, \( \alpha_s(q) \), to its value at a specific renormalization scale \( \mu \). Let us express \( \alpha_s(q) \) as a power series

\[ \alpha_s(q) = \Phi(\alpha_s(\mu), t) = \sum_{n=0}^{\infty} c_n(\mu) t^n \]

where \( t = \ln(\mu^2/q^2) \). \( \quad (4) \)
It follows that
\[ \alpha_c(\mu) = \Phi(\alpha_c(\mu), 0) = \phi_0(\mu). \] (5)

For static heavy quarks, the potential energy is a physical quantity. As a result, \( \alpha_c(q) \) obeys the renormalization-group equation
\[ \mu^2 \frac{d}{d \mu^2} \alpha_c(q) = \frac{\partial \Phi(\alpha_c, \mu)}{\partial \mu} + \beta_c(\alpha_c) \frac{\partial \Phi(\alpha_c, t)}{\partial \alpha_c} = 0, \] (6)

where the \( \beta \) function for \( \alpha_c \) is defined by
\[ \beta_c(\alpha_c) = \mu^2 \left( \frac{\partial \alpha_c(\mu)}{\partial \mu} \right) = -4\pi \sum_{n=0}^{\infty} \beta_{c,n} \left( \frac{\alpha_c(\mu)}{4\pi} \right)^{n+2}. \] (7)

Since the coefficients \( c_n(\mu) \) in Eq. (4) can be expressed as partial derivatives of \( \Phi \) with respect to \( t \),
\[ c_n(\mu) = \frac{1}{n!} \frac{\partial^n}{\partial t^n} \Phi(\alpha_c(\mu), t) \bigg|_{t=0}, \] (8)

one can show using the renormalization-group equation
\[ c_n(\mu) = \frac{1}{n!} \left( -\beta_c(\alpha_c) \frac{\partial}{\partial \alpha_c} \right)^n \alpha_c(\mu). \] (9)

Combining this expression with Eqs. (4) and (8), we find the relation between the couplings \( \alpha_c(q) \) and \( \alpha_c(\mu) \):
\[ \alpha_c(q) = \exp \left[ -t \beta_c(\alpha_c) \frac{\partial}{\partial \alpha_c} \right] \alpha_c(\mu) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left( -\beta_c(\alpha_c) \frac{\partial}{\partial \alpha_c} \right)^n \alpha_c(\mu). \] (10)

From this expression, one finds
\[ c_n(\mu) = \mathcal{S} \left( \left[ \alpha_c(\mu) \right]^{n+1} \right). \] (11)

Now we can perform the Fourier transform in Eq. (2) using the following formula:
\[ \int \frac{d^3q}{(2\pi)^3} e^{iq \cdot r} \frac{\partial^n}{\partial u^n} \mathcal{F}(r, \mu, u) \bigg|_{u=0}, \] (12)

where for \(-1 < u < 1/2,\)
\[ \mathcal{F}(r, \mu, u) = \mu^{2u} \int \frac{d^3q}{(2\pi)^3} e^{iq \cdot r}, \] (13)

with
\[ f(u) = \sqrt{\frac{\tan(\pi u)}{\pi u}} \exp \left[ \sum_{n=1}^{\infty} \frac{(2u)^{2n+1}}{2n+1} \zeta(2n+1) \right]. \]

Here,
\[ r' = r e^{-\gamma}, \] (15)
\( \gamma \equiv 0.57721 \ldots \) is Euler’s constant, and \( \zeta \) denotes the Riemann \( \zeta \) function.\(^3\) Higher order terms in the expansion of \( f(u) \) at \( u = 0 \) can be easily obtained.
using e.g. Mathematica [5]. It follows from Eqs. (12) and (13) that
\[ F(t^n) = \frac{\partial^n}{\partial u^n} \left[ e^{au} f(u) \right] \bigg|_{u=0} = \sum_{k=0}^{n} \frac{n!}{k!} f_{n-k} y^k, \]
(16)
where
\[ y = \ln (\mu^2 r^2). \]
(17)
Then from Eq. (10) we obtain
\[ \tilde{\alpha}_s(1/r) = F[\alpha_s(q)] = \sum_{n=0}^{\infty} \frac{F(t^n)}{n!} \left( -\beta_0(\alpha_s) \frac{\partial}{\partial \alpha_s} \right)^n \alpha_s(q = 1/r). \]
(18)
In particular, for \( y = 0 \), which implies \( \mu = 1/r' \),
\[ F(t^n)(y = 0) = n! f_n. \]
(19)
Hence,
\[ \tilde{\alpha}_s(1/r) = \sum_{n=0}^{\infty} f_n \left( -\beta_0(\alpha_s) \frac{\partial}{\partial \alpha_s} \right)^n \alpha_s(q = 1/r'). \]
(20)
This equation is the specific representation of the functional relation (3) which we set out to derive. The function \( \beta_0(\alpha_s) \) is defined in Eq. (7) and the coefficients \( f_n \) in Eq. (14).

Before studying consequences of Eq. (20) for the perturbative QCD potential, let us discuss first a model which we can solve numerically and which describes the energy levels of quarkonia very well. The model that we consider is the Richardson potential [6] which in momentum space reads:
\[ V_R(q) = -4\pi C_F \frac{\alpha_s^{(R)}(q)}{q^2} = -\frac{16\pi^2 C_F}{\beta_0 q^3 \ln(1 + q^2/A_R^2)}. \]
(21)
A very good description of the charmonium and bottomonium states is obtained using this potential for \( \Lambda_R = 0.4 \text{ GeV} \) and
\[ \beta_0 = 11 - \frac{2}{3} n_f, \]
(22)
where \( n_f = 3 \) is the number of the light quarks. In Fig. 1 the function \( \alpha_s^{(R)}(q) \) is plotted as the dash-dotted line for \( q \) between 1 and 10 GeV. The next-to-next-to-leading order (three-loop renormalization-group improved) QCD coupling \( \alpha_s^{(3)}(q) \) [3] for \( n_f = 4 \) active flavours and \( \alpha_{\text{MS}}(m_Z) = 0.119 \) is also shown as the solid line. 4 It is seen that in the range of \( q \) relevant for the position of I(1S) the two curves are fairly close. The QCD potential is more attractive which implies that the value of \( m_b \) extracted from the mass of I(1S) is slightly larger for the next-to-next-to-leading order QCD than for the Richardson potential. In fact in a recent article [8] \( m_b = 4.96 \text{ GeV} \) is obtained 5 with an uncertainty of about 100 MeV arising from the uncertainty in the input value of \( \alpha_{\text{MS}}(m_Z) \). For the Richardson poten-

\[ ^4 \text{We used the three-loop renormalization-group equation to evolve the } \text{MS coupling, and Eq. (9) of [7] without the } a^2 \text{-term to match the 4-flavour to the 5-flavour theory at the matching scale } 5 \text{ GeV, and } m_b(\text{pole}) = 4.88 \text{ GeV. The result is not very sensitive to varying this matching scale.} \]

\[ ^5 \text{This value is different from what [8] gives as a final value of } m_b. \text{ It corresponds to the result from the two-loop static QCD potential, which does not include a shift in } m_b \text{ of about 50 MeV due to relativistic and leading non-perturbative corrections.} \]
tial $m_s = 4.88$ GeV is obtained. All these attractive features suggest that at scales of a few GeV the model closely resembles the true QCD potential.

For this model, we study the validity of the relation (20) when it is expressed as a perturbative series of $\alpha$, in the asymptotic region where the coupling is small. In order to define the perturbative coupling of this model unambiguously in both momentum and position spaces, we subtract the confining part of the Richardson potential

$$V_R(q) = V_{\text{conf}}(q) - 4\pi C_F \frac{\alpha_R(q)}{q^2}. \quad (23)$$

The confining part $V_{\text{conf}}(q)$ corresponds to a linear confining potential in position space

$$V_{\text{conf}}(r) = \frac{2\pi C_F}{\beta_0} A_R^2 r. \quad (24)$$

The coupling

$$\alpha_R(q) = \frac{4\pi}{\beta_0} \left[ \frac{1}{\ln(1 + q^2/A_R^2)} - \frac{A_R^2}{q^2} \right] \quad (25)$$

is considered as a perturbative coupling in momentum space. The coupling $\alpha_R(q)$ is plotted as the dashed line in Fig. 1. The corresponding coupling in position space is given by the following expression [6]:

$$\alpha_R(1/r) = \frac{2\pi}{\beta_0} \left[ 1 - 4 \int_1^\infty \frac{du}{u} \frac{e^{-\alpha_R u}}{\ln^2(u^2 - 1) + \pi^2} \right] \quad (26)$$

which can be calculated numerically to a very high precision. In particular the accuracy of our numerical calculations was $10^{-10}$. From Eq. (25) we calculate the $\beta$ function

$$\beta_R(\alpha_R) = q^2 \frac{\partial \alpha_R(q)}{\partial q^2} = \frac{4\pi}{\beta_0} \alpha_R^2 + \frac{\beta_0}{4\pi} \frac{\alpha_R}{1 + q^2/A_R^2}. \quad (27)$$

where $q^2/A_R^2$ is understood as a function of $\alpha_R$. For $q \gg A_R$ the second term in the expression for $\beta_R$ is negligible. Then our problem becomes very simple: there is only one term in the sum in Eq. (7) corresponding to

$$\beta_{R,0} = \beta_0. \quad (28)$$

Instead of Eq. (20) we have a much simpler relation

$$\tilde{\alpha}_R(1/r) \sim \sum_{n=0}^\infty \alpha_R(q) \left( \frac{\beta_0 \alpha_R(q)}{4\pi} \right)^n n! f_n, \quad (29)$$

with $q = 1/r$. The symbol ‘‘$\sim$’’ indicates that in our derivation non-perturbative higher twist contributions to $\beta_R$ (terms suppressed by powers of $e^{-4\pi/(\beta_0 \alpha_R^2)}$) have been neglected. Strictly speaking, the relation (29) may be valid only for $q \gg A_R$ i.e. for small $\alpha_R$. We have computed the first twelve coefficients $f_n$ and checked that all are positive. For a given $1/r$, the terms in (29) decrease with increasing $n$ for $n < n_{\text{min}}$ and increase for $n > n_{\text{min}}$. The value $n_{\text{min}}$ corresponding to the minimal contribution increases with growing $1/r$. The best approximation is obtained when the series in (29) is truncated at $N = n_{\text{min}}$. In Table 1 the values are given of the functions $\tilde{\alpha}_R(1/r)$ computed numerically from Eq. (26) and $\tilde{\alpha}_{R,N}(1/r)$ obtained from the asymptotic expansion (29) truncated at $n = N$. The coefficient $\beta_0$ is evaluated for $n_f = 3$ and $1/r$ is varied between 10 and $10^7$ GeV. It is evident that for large values of $1/r$ a very good approximation is obtained using the truncated series $\tilde{\alpha}_{R,N}(1/r)$. However, at $1/r = 10$ GeV, $n_{\text{min}} = 1$ and the quality of approximation practically does not improve when instead of

$$\tilde{\alpha}_R(1/r) = \alpha_R(q = 1/r) \quad (30)$$

the formula including the $\sigma(\alpha_R^2)$ term is used. At $1/r = 20$ GeV inclusion of the cubic term improves the quality of approximation, and at $1/r = 100$ GeV also the term $\sigma(\alpha_R^2)$ is needed. Let the truncation at $n = n_+$ result in the closest approximation of $\tilde{\alpha}_R$ from above and $\delta \tilde{V}_s(r)$ denote the difference between the exact and the approximate values of the potential $\tilde{V}_s(r)$. If the contribution of the term for $n = n_+$ is not included, i.e. the series in (29) is truncated at $n = n_+ - 1$ the closest approximation from below is obtained. Let $\delta \tilde{V}_s(r)$ be the corresponding difference of the potentials. The values of
the functions $\delta \overline{V}_r(r)$ are also given in Table 1. One can estimate the theoretical uncertainty of the perturbative approach by considering the change in the value of the potential

$$\delta \overline{V}(r) = \delta \overline{V}_+ - \delta \overline{V}_-,$$

(31)
corresponding to truncating the asymptotic formula at $n = n_+ - 1$ and $n = n_+$. It is interesting that $\delta \overline{V}(r)$ does not change drastically with $r$ and is of the order of $A_R$.

In perturbative calculations of the energy levels of a $\bar{Q}Q$ system in this model, the truncation of the asymptotic series for the potential leads to an uncertainty in the perturbative pole mass $m_\omega$ of the heavy quark. This uncertainty is of the order of $\delta m \sim \delta \overline{V}(r_0)/2$, where $r_0$ denotes the typical size of the bound-states. This estimation is obtained assuming that the shift in the binding energy due to the change of the coupling $\pi_R$ is equal to $-\delta \overline{V}(r_0)$, which would be the case for the Coulomb potential. The shift in the binding energy is then compensated by shifting the masses of $Q$ and $\overline{Q}$ by $\delta m$. One may think that for a stable quark $Q$ the uncertainty in $m_\omega$ can be reduced by comparing the $\bar{Q}Q$ bound states of different sizes, thereby disentangling the correlation in the dependences on the coupling and mass. However, as it has already been mentioned the variation of $\delta \overline{V}(r)$ with $r$ is quite moderate. Moreover at larger distances non-perturbative effects become important. So, in fact it is difficult to reduce the uncertainty in the mass determination significantly. 7

It may appear too pessimistic to estimate $\delta m$ from $\delta \overline{V}$. One can argue that instead of considering the truncated asymptotic series for $\alpha_R$ the exact expression (26) should be used in the Schrödinger equation for the $\bar{Q}Q$ system. Then $m_\omega$ can be determined from the mass of the $\bar{Q}Q$ bound-states without any theoretical uncertainty. This argument is based, however, on the complete knowledge of the function $\alpha_R(q)$ including the range of its argument where non-perturbative contributions are very important. In purely perturbative calculations $\delta m$ has to be interpreted as a perturbative theoretical uncertainty in $m_\omega$ due to the asymptotic character of the series (29).

It seems reasonable to expect that the qualitative features of the asymptotic expansion (20) should be similar for the Richardson potential and for the QCD potential. This is a useful assumption because only

---

### Table 1

Comparison of the coupling $\pi_R(1/r)$ for the Richardson potential and the best approximation obtained by truncating the asymptotic series (29) at $n = N$

<table>
<thead>
<tr>
<th>$1/r$ [GeV]</th>
<th>$N$</th>
<th>$\pi_R(1/r)$</th>
<th>$\pi_R(1/r)$</th>
<th>$\delta \overline{V}$ [GeV]</th>
<th>$\delta \overline{V}$ [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>3</td>
<td>0.2856</td>
<td>0.2714</td>
<td>-0.193</td>
<td>0.189</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>0.22252</td>
<td>0.22183</td>
<td>-0.381</td>
<td>0.021</td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td>0.17647</td>
<td>0.17470</td>
<td>-0.222</td>
<td>0.118</td>
</tr>
<tr>
<td>$10^2$</td>
<td>4</td>
<td>0.14868</td>
<td>0.14912</td>
<td>-0.059</td>
<td>0.183</td>
</tr>
<tr>
<td>$10^3$</td>
<td>6</td>
<td>0.098937</td>
<td>0.099017</td>
<td>-0.106</td>
<td>0.120</td>
</tr>
<tr>
<td>$10^4$</td>
<td>9</td>
<td>0.07413536</td>
<td>0.07413084</td>
<td>-0.143</td>
<td>0.060</td>
</tr>
<tr>
<td>$10^5$</td>
<td>11</td>
<td>0.05938222</td>
<td>0.05938213</td>
<td>-0.175</td>
<td>0.012</td>
</tr>
<tr>
<td>$10^6$</td>
<td>12</td>
<td>0.04958339</td>
<td>0.04958315</td>
<td>-0.409</td>
<td>0.318</td>
</tr>
<tr>
<td>$10^7$</td>
<td>12</td>
<td>0.04257994</td>
<td>0.04257992</td>
<td>-1.008</td>
<td>0.185</td>
</tr>
</tbody>
</table>

---

6 This can be shown easily for a small $\alpha_R$. Since asymptotically $f_n \sim 2^n$, it follows that $n_{min} \sim 2 \pi / (\beta_0 \alpha_R) \sim \ln(1/(\alpha_R r))$. Then one can estimate the size of the last term included or rejected in Eq. (29) to be $\delta \pi_R \sim A_R r$, which leads to $\delta \overline{V}(r) \sim A_R$.  

7 Behind this argument for the Richardson potential lies a following corresponding perspective in QCD. In principle one is able to determine the two independent fundamental parameters of QCD, $m_\omega$ and $\alpha_s$, from a sufficient number of different physical observables $\mathcal{O}(m_\omega, \alpha_s)$ by exploiting their different dependences on these parameters. When we express the observables in perturbative expansions of $\alpha_s$, however, we encounter uncertainties of the order of $A_{QCD} \sim 200$ MeV originating from truncations of the asymptotic series. Then, if we try to extract the values of $m_\omega$ and $\alpha_s$ from these observables, the uncertainties should be attached to both of these parameters which cannot be reduced within a purely perturbative approach.
limited information is available in the latter case. In perturbative QCD the coupling $\alpha_s$ is expressed as an asymptotic series in $\alpha_{\text{MS}}$ and only the first three terms are known [2,3]. In Fig. 2, $\alpha_s(q)$ is shown for $n_f = 5$ active flavours and $q$ between 10 and 100 GeV. The solid line corresponds to $\alpha_{\text{MS}}(m_Z) = 0.118$ and the dashed lines to changing this input value by $\pm 0.003$ which is the present error of this parameter [9]. In QCD the first three coefficients of the $\beta$ function (7) are also known [10,2,3]:

$$\beta_{V,0} = \beta_0,$$

$$\beta_{V,1} = \beta_1 = 102 - \frac{38}{\pi} n_f,$$

$$\beta_{V,2} = 1854 - \frac{2239}{6} n_f + \frac{377}{24} n_f^2$$

$$+ \pi^2 \left( 18 - \frac{3 \pi^2}{4} \right) (33 - 2 n_f)$$

$$+ \left( 726 - \frac{204}{\pi} n_f + \frac{94}{\pi n_f^2} \right) \xi (3).$$ (32)

Using this information one can write a formula for $\alpha_s(1/r)$ which is accurate up to fifth order in $\alpha_s$:

$$\alpha_s = \alpha_s \left( 1 + 2 \beta_0 \frac{r}{x} a_0^s + 5 \beta_0 \beta_1 f_2 + 6 \beta_0^3 f_3 \right) a_1^s$$

$$\left[ 3 \beta_0^2 f_2 + 6 \beta_0 \beta_{V,2} f_2 + 26 \beta_0^3 \beta_1 f_3 + 24 \beta_0^2 f_4 \right] a_2^s + \mathcal{O}(\alpha_s^3).$$ (33)

where $\alpha_s(1/r) = \alpha_s(q = 1/r')$ and $a_s = \alpha_s/(4 \pi)$. Let us remark that the term quadratic in $\alpha_s$ is absent because it has been absorbed by shifting the argument of the coupling $\alpha_s$. After simple algebra the following formula is derived:

$$\alpha_s = \alpha_s \left( 1 + 11 - \frac{2}{3} n_f \right) \frac{\pi^2}{3} a_1^s$$

$$+ \left[ \left( 935 - \frac{1555}{3} n_f + \frac{190}{3} n_f^2 \right) \pi^2 \right.$$$$

$$+ \left. \left( 21296 - 3872 n_f - \frac{704}{3} n_f^2 - \frac{128}{\pi} n_f^3 \right) \xi (3) \right] a_2^s$$

$$+ \left( \left( 25596 - \frac{3797}{6} n_f + \frac{21913}{34} n_f^2 - \frac{377}{91} n_f^3 \right) \pi^2 \right.$$$$

$$- \frac{6633}{405} n_f^3 + \frac{304}{1275} n_f^4 \right) \pi^4$$

$$+ \left( - \frac{1029}{4} + 33 n_f - n_f^2 \right) \pi^6$$

$$+ \left( 855712 - \frac{1889888}{9} n_f \right.$$$$

$$+ \frac{432640}{24} n_f^2 - \frac{3616}{81} n_f^3 \right) \xi (3)$$

$$+ \left( 7986 - \frac{9196}{9} n_f + \frac{2552}{n_f^2} n_f^3 \right. - \frac{208}{27} n_f^4 \right) \pi^2 \xi (3) a_3^s$$

$$+ \mathcal{O}(\alpha_s^4).$$ (34)

The numerical values of the coefficients in the above expansions for $n_f = 5$ are equal to:

$$\alpha_s = \alpha_s + 1.22454 \alpha_s^3 + 5.59618 \alpha_s^4 + 32.2015 \alpha_s^5$$

$$+ \mathcal{O}(\alpha_s^6).$$ (35)

whereas for $n_f = 4$ and $n_f = 3$ one obtains

$$\alpha_s = \alpha_s + 1.44676 \alpha_s^3 + 7.38182 \alpha_s^4 + 46.3717 \alpha_s^5$$

$$+ \mathcal{O}(\alpha_s^6).$$ (36)

$$\alpha_s = \alpha_s + 1.68750 \alpha_s^3 + 9.45282 \alpha_s^4 + 64.0389 \alpha_s^5$$

$$+ \mathcal{O}(\alpha_s^6).$$ (37)

respectively. The numerical coefficients in the relations (35)–(37) are large. The asymptotic character of the expansion is evident and only for very short distances (1/r ≥ 30 GeV) all five terms should be kept. At such distances the number of active flavours is $n_f = 5$. The $\mathcal{O}(\alpha_s^4)$ and $\mathcal{O}(\alpha_s^5)$ terms in (35) are equal for $\alpha_s = 0.1738$ and their contributions to $\alpha_s$ are 0.0051. This implies that the truncation of the
Asymptotic series leads to an uncertainty of about 2.8% in $\bar{\alpha}_s$. The value of $\alpha_s$ corresponds to $1/r = 32$ GeV, a distance which is not very different from those probed in $t \bar{t}$ production near the threshold. Phenomenological consequences of the above observation are discussed elsewhere [11]. At the distances probed by $b \bar{b}$ states, the corresponding values of $\alpha_s$ are so large that the quartic term in the expansion (34) must be rejected. Thus it would be simply inconsistent to keep other contributions of this order like those in the relation between $\alpha_s$ and $\alpha_{\overline{MS}}$ even if this relation were known. Our analysis leads to a rather surprising conclusion that in the framework of purely perturbative QCD, theoretical uncertainties cannot be reduced below the present level! (This is consistent with the discussion given in [8].) If we accept such a radical point of view, we can proceed even further by observing that the quartic term in (34) depends only on $\beta_0$ and is the same for the QCD static potential and for the Richardson potential. The latter case suggests that it is not obvious at all if the inclusion of the cubic term improves the accuracy of the formula for $\bar{\alpha}_s$. It may well be that a more precise answer is obtained if the cubic term is also rejected. Without extra information beyond the purely perturbative approach, we cannot answer the question: to reject or not to reject. Clearly, keeping the cubic term implies a stronger attraction between the quarks. Thus, if $m_b$ is determined from the energy of the $b \bar{b}$ ground state, the value of $m_b$ obtained using a perturbative calculation and the potential including the cubic term is larger than the value corresponding to rejecting this term. The difference in $m_b$ is of the order of 100 MeV and our analysis indicates that attempts to reduce this error to a much smaller value within perturbative QCD may be inconsistent.

Acknowledgements

This work is partly supported by KBN grant 2P03B08414, by BMBF grant POL-239-96 and by the Alexander von Humboldt Foundation.

References

Time ordering in off-diagonal parton distributions

Markus Diehl a, Thierry Gousset b

a CPT, Ecole Polytechnique, 91128 Palaiseau, France
b NIKHEF, P.O. Box 41882, 1009 DB Amsterdam, The Netherlands

Received 22 January 1998
Editor: R. Gatto

Abstract

We investigate the relevance of time ordering in the definition of off-diagonal parton distributions in terms of products of fields. The method we use easily allows determination of their support properties and provides a link to their interpretation from a parton point of view. It can also readily be applied to meson distribution amplitudes. © 1998 Elsevier Science B.V. All rights reserved.

1. Recently there has been renewed interest in off-diagonal parton distributions, i.e. in correlation functions of quark or gluon fields between different nucleon states, which have been introduced some time ago [1]. They appear for instance in the description of exclusive photon or meson production in $\gamma^* p$ collisions at large photon virtuality $Q^2$ and small squared momentum transfer $t$ of the proton [2–5], showing up as the long distance nonperturbative quantities when one factorises the transition amplitude into short and long distance sub-processes. In meson electroproduction there is yet another nonperturbative ingredient, namely the meson distribution amplitude which describes the transition from the valence quarks to the final vector meson. The corresponding diagrams are shown in Fig. 1.

The relevance of off-diagonal parton distributions is not restricted to these processes. In particular the asymmetric gluon distribution has been considered in several $\gamma^* p$ processes at small Bjorken $x$, such as diffractive production of dijets and of heavy quarks [6]. For exclusive $J/\psi$ production one may relax the requirement of large $Q^2$ because of the large meson mass [7], as well as for the exclusive production of a $Z$ boson [8].
2. Let us first have a closer look at the kinematics of the reactions we have in mind. We denote particle momenta according to
\begin{equation}
\gamma^+(q) + p(p) \rightarrow A(q') + p'(p') ,
\end{equation}
where \( A \) is a meson, real photon or heavy gauge boson, or a quark-antiquark pair in a diffractive process. One might also have a virtual \( Z \) or \( W \) instead of the \( g \) and \( r \) or \( A \) may be charged, replacing the initial or final state proton by a neutron as necessary. We will use the momentum transfer \( \Delta = p - p' \) from the proton, and the invariants \( t = \Delta^2, W^2 = (p + q)^2, Q^2 = -q^2, m_A^2 = q^2, m_p^2 = p^2 \).

We now consider those reference frames where the incoming photon and proton are collinear. Picking out any one of them we choose the \( z \) axis along the proton momentum \( p \) and introduce two lightlike vectors
\begin{equation}
v = \frac{1}{\sqrt{2}} (1,0,0,1), \quad v' = \frac{1}{\sqrt{2}} (1,0,0,-1) ,
\end{equation}
which respectively set a plus and minus direction. We can write
\begin{equation}
p^\mu = p^+ v^\mu + \frac{m_p^2}{2 p^+} v'^\mu, \quad \Delta^\mu = \xi p^+ v^\mu - \frac{m_p^2 \xi + \Delta^2}{2 p^+ (1 - \xi)} v'^\mu + \Delta^\mu
\end{equation}
and
\begin{equation}
t = - \frac{m_p^2 \xi^2 + \Delta^2}{1 - \xi} ,
\end{equation}
having defined the variable \( \xi = \Delta^+ / p^+ \), whose exact expression in terms of the invariants listed above we will not need here. We further have
\begin{equation}
q^\mu = - x_N p^+ v^\mu + \frac{Q^2}{2 p^+ x_N} v'^\mu ,
\end{equation}
where
\begin{equation}
x_N = \frac{2 x_B}{1 + \sqrt{1 + 4 x_B^2 m_p^2 / Q^2}}
\end{equation}
is Nachtmann’s and \( x_B = Q^2 / (2 p \cdot q) \) Bjorken’s scaling variable. From the positivity of particle energies in the final state one obtains
\begin{equation}
x_N \leq \xi < 1 .
\end{equation}

To ensure that the blob \( H \) in Fig. 1 corresponds to a hard scattering we require that at least one of \( Q^2 \) and \( m_A^2 \) be large compared with a GeV$^2$. For the blob \( S \) to be a soft, long distance quantity one will need that the transverse bend \( |\Delta_T| \) is of the order of a hadronic scale.
The threshold region is special in its dynamics: apart from the formation of resonances one can expect strong rescattering effects between $A$ and the outgoing proton which may destroy factorisation. One will ask for $W - m_A - m_q \gg 1$ GeV to exclude this region.\(^5\) It can be shown that this is equivalent to limiting the range of $\xi$ to $1 - \xi \gg (m_q W)/(W^2 + Q^2)$, which according to (4) also puts an upper limit on $-t$. Up to corrections of order $m_q^2/(W^2 - m_A^2)$ and $\Delta q^2/(W^2 - m_A^2)$ one then has

$$\xi = \frac{Q^2 + m_A^2}{2 p \cdot q}. \quad (8)$$

Under these conditions one can find a frame, e.g. the $\gamma^* p$ c.m., where the initial proton is moving fast and collides head on with the photon, $p^+ \gg m_p$, and where the scattered proton is fast as well, $p'^+ \gg m_p$. Such a frame is natural for a partonic interpretation of our process.\(^6\)

3. The nonperturbative transition from the proton to the parton level (see Fig. 1) is described by a two-point function of the form

$$M = \frac{1}{2\pi} \int dz^w e^{i x p \cdot z} \langle p', \sigma' \mid T \bar{\psi}(0) \gamma^+ \psi(z) \mid p, \sigma \rangle \bigg|_{\mu = \epsilon = \epsilon'}. \quad (9)$$

This representation holds in the $A^+ = 0$ gauge, where $A^\mu$ is the gluon potential; for other gauges one has to insert the standard path ordered exponential between the two quark fields. Instead of $\gamma^+$ in (9) there can be other Dirac matrices, which in order to give a leading twist contribution must transform like $\gamma^+$ under a longitudinal boost. Which matrices are relevant depends on the process considered; in virtual Compton scattering, $\gamma^+ p \to \gamma p$, one needs for instance $\gamma^+$ and $\gamma^+ \gamma_5$. There are also contributions from two-point functions where gluon field strengths replace the quark fields in (9).

We notice that these quark and gluon two-point functions depend on the spins $\sigma$ and $\sigma'$ of the initial and final state proton. This spin structure can be made explicit by writing (9) as a sum of Dirac bilinears for the proton multiplied by scalar functions $F(x, \xi, t)$. It is these functions that are called off-diagonal parton distributions; we will not need their explicit definitions here.\(^7\) For ease of writing we will not explicitly display the labels $\sigma$ and $\sigma'$ henceforth.

The usual parton distributions that occur in inclusive deep inelastic scattering involve diagonal matrix elements of quark fields or gluon field strengths, i.e. they correspond to the limit $p' = p$ and $\sigma' = \sigma$ in (9). The variables $\xi = \Delta^+/p^+$ and $t = \Delta^2$ are then zero since $\Delta = 0$. The diagrams that describe inclusive deep inelastic scattering are obtained by cutting the ones of Fig. 1(a) with the real photon replaced by a virtual one.

At this point we note that because the proton states in (9) are different the off-diagonal distributions do not have any probabilistic interpretation, contrary to the usual parton distributions. In particular one cannot expect the $F_i$ to have a definite sign as ordinary quark or gluon densities, which are nonnegative for $0 < x < 1$.

4. When calculating the transition amplitude from the diagrams in Fig. 1 the soft blob describing the proton-parton coupling translates into the Fourier transformed matrix element of a time ordered product of parton operators. We will show that one can actually drop the time ordering and instead use ordinary products.

The same problem has long ago been considered by Jaffe\(^9\) for the parton distributions in deep inelastic scattering. For the deep inelastic cross section one needs the absorptive part of the forward $\gamma^+ p \to \gamma^+ p$ amplitude, cutting the diagrams including the soft blob, and naturally obtains ordinary products instead of time

\(^5\) We remark that we will not actually make use of this condition in the arguments of this paper.

\(^6\) The analysis we perform in this paper may also be done in frames where $p$ and $q$ have nonzero transverse components\(^2\), provided that these are sufficiently small, as $\Delta f$ in the frames considered here.

\(^7\) A comparison of the various distributions introduced in the recent literature can be found in [5].
ordered ones. It was however shown in [9] that the ordinary product already appears in the full amplitude, not only its absorptive part. The same result also follows from early work by Landshoff and Polkinghorne [10], without being explicitly stated there, and we will apply their method to the non-diagonal case here.

The fact that one can omit time ordering in the parton distributions has important consequences for their support properties in the scaling variable \(x\), which in turn are crucial for their interpretation in terms of partons [9,11]. In the off-diagonal case it is also needed if one wants to derive a dispersion relation for the scattering amplitude [4]. Furthermore it allows one to constrain the two-point function (9) using time reversal invariance: together with parity invariance one finds that the phases of the scalar parton distributions \(F_j(x,\xi,t)\) are a matter of convention, so that they can be defined as real valued.

As a by-product of the demonstration we will obtain the support properties for \(F_j\) in \(x\) and identify different regions in \(x\) according to whether the partons are in the initial or final state of the hard scattering \(H\). This generalises the results obtained in [9] concerning the support and partonic interpretation of the usual densities to the case of off-diagonal distributions.

5. Let us first consider the quark two-point function \(M\) of (9). Our arguments will be unchanged if \(\gamma^+\) is replaced by another suitable Dirac matrix. We rewrite

\[
M = \frac{1}{(2\pi)^2} \int d^2k^+ d^2k^- \mathcal{A}[k^+=x \gamma^+],
\]

where

\[
\mathcal{A} = \int d^4z e^{ik^+z} \langle p' | T\overline{\psi}(0) \gamma^+ \psi(z) | p \rangle
\]

is an amplitude describing the scattering of an off-shell antiquark with momentum \(-k\) on an on-shell proton with momentum \(p\),

\[
p(p) + \overline{q}(-k) \rightarrow p(p') + \overline{q}(-k').
\]

Note that this amplitude is not truncated in the parton legs. The plus components \(k^+ = xp^+\) and \(k'^+ = x'p^+\) with \(x' = x - \xi\) are kept fixed in the following.

The kinematical invariants \(\mathcal{A}\) depends on are the virtualities \(k^2\) and \(k'^2\) and the Mandelstam variables \(s = (-k + p)^2\), \(u = (k + p')^2\) and \(t\), the latter being fixed by the kinematics of the reaction (1). Following [10] we assume that the analytical properties of \(\mathcal{A}\) are the usual ones known from the study of Feynman diagrams, i.e. that it has cuts for nonnegative \(\text{Re} \ s, \text{Re} \ u\) and singularities for nonnegative \(\text{Re} \ k^2\) and \(\text{Re} \ k'^2\). In the standard conventions, these singularities are slightly below the real axis of these variables.

We can carry out the integration over \(k^-\) in (10) taking into account the singularity structure of \(\mathcal{A}\). To see

---

8 This is also plausible from the point of view of the operator product expansion. There one obtains matrix elements of local operators, which are moments of the parton distributions; only the Wilson coefficients describing the hard scattering depend on whether one takes the full amplitude or its imaginary part.

9 The support of \(F_j\) has been obtained in [5] by analysing the soft blob \(S\) in terms of Feynman graphs.

10 To avoid possible complications due to the dependence of \(\mathcal{A}\) on the spin of the proton one can apply our argument directly to the scalar functions \(F_j(x,\xi,t)\) where the proton spin structure has been factored off. The corresponding amplitudes for (12) only depend on the above invariants and on the plus components \(p^+, \Delta^+\) and \(k^+\) which are all fixed. The dependence on plus components comes from factoring off the proton spin and from the Dirac matrix between the quark fields, and also from the choice of gauge \(A^+ = 0\).
where the singularities are located in the complex $k^{-}$ plane we in turn express $k^{-}$ through one of the invariants $s, k^{2}, k^{'2}, u$, supplemented by $x, k_{T}$ and the variables in (3) which are all kept fixed during the $k^{-}$ integration:

$$k^{-} = \frac{s + k_{T}^{2}}{2p^{+}(x - 1)} + p^{-}, \quad k^{-} = \frac{k^{2} + (k_{T} - \Delta_{T})^{2}}{2p^{+}(x - \xi)} + \Delta^{-}, \quad k^{-} = \frac{k^{'2} + k_{T}^{2}}{2p^{+}x},$$

$$k^{-} = \frac{u + (k_{T} - \Delta_{T})^{2}}{2p^{+}(x + 1 - \xi)} - p^{-} + \Delta^{-}.$$  \hspace{1cm} (13)

There is a sequence of sign changes of the above denominators as $x$ varies from $-\infty$ to $+\infty$. Using $0 < \xi < 1$ we see that they occur at $x = \xi - 1, 0, \xi$ and 1. These signs determine whether the singularities of $\mathcal{M}$ are located above or below the real axis in the $k^{-}$ plane.

For $x \leq \xi - 1$ and $x \geq 1$ all singularities lie on the same side of the real $k^{-}$ axis and we can close the contour in the other half plane without encircling any singularity. The $k^{-}$ integral vanishes and we deduce that $M$ is only nonzero for $\xi - 1 < x < 1$.

To be able to close the integration contour in the $k^{-}$ plane we have made the dynamical assumption that $\mathcal{M}$ vanishes sufficiently fast as $k^{-}$ tends to infinity in the complex plane while $k^{+}$ and $k_{T}$ are fixed, a limit in which the parton virtualities $k^{2}$ and $k^{2}$ become infinite. Note that we have not yet integrated over $k_{T}$; as remarked in [9] the integration of $\mathcal{M}$ over $k_{T}$ and $k^{-}$ leads to ultraviolet divergences, but this does not contradict our assumption that the integral over $k^{-}$ alone is sufficiently well behaved. The full integration over $k_{T}$ and $k^{-}$ can only be done in a suitably regularised theory; it is known from the usual parton distributions that the corresponding renormalisation of the integral (10) is intimately related with their logarithmic evolution in QCD.

Let us now discuss in turn the regions of $x$ in the interval $\xi - 1 < x < 1$. Details of the argument are given in the Appendix.

1. In the region $\xi \leq x < 1$, the $s$ singularities are in the upper half plane of $k^{-}$ whereas all other singularities are in the lower half plane. Wrapping the $k^{-}$ contour around the $s$ cuts gives

$$\int_{-\infty}^{+\infty} dk^{-} \mathcal{M} = \int dk^{-} \text{disc}_{s} \mathcal{M},$$

where $\text{disc}_{s} \mathcal{M}$ is the discontinuity of $\mathcal{M}$ across the cuts in $s$. We use now the standard reasoning on the matrix element

$$\hat{M} = \frac{1}{2\pi} \int d^{2}k_{T} dk^{-} e^{ik_{T}p^{+}} \left\langle p' | \bar{\psi}(0) \gamma^{+} \psi(z) | p \right\rangle |_{z = z'},$$

$$\hat{M} = \frac{1}{(2\pi)^{2}} \int d^{2}k_{T} dk^{-} \int d^{2}z e^{ik_{T}p^{+}} \left\langle p' | \bar{\psi}(0) \gamma^{+} \psi(z) | p \right\rangle |_{z = x, \Delta},$$

which does not involve time ordering, inserting a complete set of intermediate states between $\bar{\psi}(0)\gamma^{+}$ and $\psi(z)$ as illustrated in Fig. 2(a). In line with our assumptions on its analyticity properties we assume that one can apply the cutting rules to $\mathcal{M}$ as one would for Feynman diagrams, and obtain

$$M = \frac{1}{(2\pi)^{2}} \int d^{2}k_{T} dk^{-} \text{disc}_{s} \mathcal{M} = \hat{M}.$$  \hspace{1cm} (16)

In this region of $x$ we can interpret $M$ as the amplitude for the proton emitting a quark with light cone fraction $x$ and reabsorbing it with light cone fraction $x'$. 


2. In the region $\xi - 1 < x \leq 0$, the $u$ singularities are alone below the real $k^-$ axis and we can write
\[
\int_{-\infty}^{+\infty} dk^- \mathcal{A} = \int dk^- \text{disc}_s \mathcal{A}.
\]

We now consider the scattering process
\[
p(p) + q(k') \rightarrow p(p') + q(k),
\]
whose amplitude is given by $-\mathcal{A}$, where the minus sign comes from interchanging the quark and antiquark field under the time ordering in (11). Defining
\[
\hat{M}' = -\frac{1}{2\pi} \int d^2 k_e dk^- \text{disc}_u \mathcal{A},
\]
and repeating our above argument we find
\[
M = \frac{1}{(2\pi)^4} \int d^2 k_e dk^- \text{disc}_u \mathcal{A} = \hat{M}'.
\]

In the present region of $x$ we can interpret $\hat{M}$ as corresponding to the emission by the proton of an antiquark with momentum fraction $-x'$ and its reabsorption with momentum fraction $-x$.

3. The intermediate region, $0 < x < \xi$, is more involved since the singularities in $s$ and $k^2$ lie in the upper half plane of $k^-$ while those in $u$ and $k^2$ lie in the lower half plane. We notice that in this case $x > 0$ and $x' < 0$, i.e. in this situation one has a quark and an antiquark flowing out of the initial state. This configuration is specific of the off-diagonal matrix element, and in [3,5] its reminiscence of the quark-antiquark distribution amplitude of a meson has been emphasised.

If we choose to close the $k^-$ contour in the upper half plane we can write the integral over $k^-$ of $\mathcal{A}$ as a sum of terms due to the singularities in $s$ and in $k^2$,
\[
\int_{-\infty}^{+\infty} dk^- \mathcal{A} = \int dk^- \left( \text{disc}_s \mathcal{A} + \{\text{terms from } k^2 \text{ singularities}\} \right),
\]
where the singularities in $k^2$ are cuts and a mass pole, cf. the Appendix. Having chosen this representation we consider again $\hat{M}$ and the scattering (12), but now when we insert a complete set of states, the kinematics allow intermediate states that already contain the outgoing proton, $|XX\rangle = |p'X'\rangle$. For those states there are diagrams which are disconnected at the right-hand side of the cut, cf. Fig. 2(b), and correspond to a cut in the variable $k^2$. They give just the extra terms in (21) and the cutting rules now allow us to write
\[
M = \frac{1}{(2\pi)^4} \int d^2 k_e dk^- \left( \text{disc}_s \mathcal{A} + \{\text{terms from } k^2 \text{ singularities}\} \right) = \hat{M}'.
\]

In all three regions of $x$ there are of course two ways to pick up singularities in the $k^-$ plane: those in the upper half will lead to $\hat{M}$ and those in the lower half to $\hat{M}'$. The situation is summarised in Fig. 2. In regions A and B we have a simple physical interpretation if we pick up the $s$ and $u$ cut, respectively, as was already discussed for diagonal densities in [9], whereas in region C the choices $s, k^2$ and $u, k^2$ appear as symmetric to each other.

For all values of $x$ we then have
\[
M = \hat{M} = \hat{M}',
\]
i.e. we can drop the time ordering in $M$ and write quark and antiquark field in any order. Our argument shows that the origin of this is the integration over $k^-$, which corresponds to fields being evaluated at the same light

\[\text{footnote}{\text{The particular singularity structure of } \mathcal{A} \text{ in this region of } x \text{ has also been remarked in [5,12].}}\]
cone time \((z^+ = 0)\). We could in fact apply it without integrating over \(k_T\) at all to obtain the analogue of (23) for \(k_T\) unintegrated parton densities.

6. Our preceding discussion can be applied to off-diagonal gluon distributions in an analogous way by considering the amplitude for the scattering of an off-shell gluon on the proton. Note that one then deals with products of gluon potentials \(A^\mu\), while the gluon distributions are defined in terms of gluon field strengths \(G^{\mu\nu}\). The passage from one to the other is however simple in the \(A^+ = 0\) gauge, since the leading twist distributions only involve \(G^{++}\) which in this gauge reduces to \(\partial^+ A^+\). For a further discussion we refer to [5].

In the case of gluons there is no minus sign when one interchanges the order of the fields in the amplitude (11) as there was in point B. above, and correspondingly the analogue of \(M'\) is defined without a minus sign.

We also remark that there are symmetry relations for the gluon distributions under the change \(x \rightarrow -x' = \xi - x\) in their first argument, so that it is enough to know them for \(x \geq \xi/2\). In fact the diagrams in Fig. 1 with gluon lines between \(S\) and \(H\) remain the same if one interchanges the light cone fractions \(x\) and \(-x'\), since \(H\) denotes a sum of several Feynman graphs.

7. As mentioned in the beginning, the diagrams of Fig. 1(b) for exclusive meson production contain not only off-diagonal parton distributions as nonperturbative input, but also the meson distribution amplitude, which is expressed in terms of a matrix element between the vacuum and the meson state \(|M\rangle\). It is easy to show along the lines of argument developed so far that the time ordering of quark operators in the distribution amplitude can be omitted, just as in parton distributions, as we will briefly outline.
The distribution amplitude is defined in terms of a two-point function
\[
\frac{1}{2\pi} \int dz e^{-i q^\mu z^\mu} \left( M | T \bar{\psi}(z) \gamma^\mu \psi(0) \right) | 0 \rangle_{\mu = z^\mu,}
\]
(24)
or a corresponding one with a different Dirac matrix; in order to suppress the path ordered exponential between the quark fields one now needs the $A^- = 0$ gauge. In analogy to (10) this two-point function can be rewritten as an integral over $l_T$ and $l^+$ at fixed $l^+ = y q^-$ of the amplitude for the transition
\[
q(l) + \bar{q}(l') \rightarrow M(q')
\]
from the valence quarks to the meson. This amplitude depends on the invariants $l^2$ and $l^+_s$. Writing
\[
l^+ = \frac{l^2 + l_T^2}{2 q^- y}, \quad l^+ = \frac{l^2 + (l_T - \Delta_T)^2}{2 q^- (y - 1)} + q^+-
\]
and closing the contour of the $l^+$ integration in the complex plane one immediately obtains that the distribution amplitude vanishes outside the region $0 < y < 1$. There one can express it in terms of singularities in either $l^2$ or $l^+_s$, obtaining matrix elements of $\bar{\psi}(z) \gamma^\mu \psi(0)$ or $-\psi(0) \bar{\psi}(z) \gamma^\mu$, respectively. The two possibilities are shown in Fig. 3.

Comparing with the present situation we see again the hybrid nature of off-diagonal parton distributions in the region $0 < x < \xi$: the cuts in $s$ or $u$ are as in ordinary, diagonal parton distributions, whereas the singularities in the parton virtualities $l^2$ or $l^+_s$ are reminiscent of distribution amplitudes.

8. The fact that parton distributions and meson distribution amplitudes can be expressed in terms of cut amplitudes is important if one wants to show that the $s$-channel discontinuities of the $\gamma^* p$ amplitudes shown in Fig. 1 are obtained from appropriate cuts of the hard blob $H$, the soft quantities being "already cut". Already of interest in the case of inclusive deep inelastic scattering this is essential for processes involving off-diagonal parton distributions, for instance when establishing a dispersion relation [4]. We shall not go into details here but wish to make some observations in connection with the results we have obtained so far.

There are different ways to cut the diagrams of Fig. 1 in the overall $s$-channel, i.e. in the variable $(p + q)^2$. If the hard blob $H$ is cut in $(q + k)^2$, which requires $x_p \leq x \leq 1$ to ensure positive energy across the cut, one can cut the soft blob $S$ in $s = (p - k)^2$; if $x_p \leq \xi$ one can also cut $S$ in $k^2$ or cut the parton line $k'$ itself in the diagrams. The two-point function $M$ can be written in terms of just these cuts. As discussed in the appendix it includes in particular the term that comes from cutting the parton line $k'$ and leads to a mass pole in the amplitude $\mathcal{A}$ of (11); hence this mass pole does not appear as an extra term in the expression of the discontinuity of the $\gamma^* p$ amplitude.

![Fig. 3. The two possibilities to pick up singularities in the $l^+$ plane for a meson distribution amplitude in the region $0 < y < 1$.](image-url)
$H$ can also be cut in $(q-k')^2$, then one must have $\xi-1 \leq x \leq \xi-x_N$. This comes with cuts of $S$ in $u=(p+k')^2$, and if $x \geq 0$ also with cuts in $k^2$ as well as the pole from cutting the parton line $k$. In this situation one will use the representation of $M$ in terms of singularities in $u$ and $k^2$. It is interesting to note that if there is a region $x_N \leq x \leq \xi-x_N$, e.g. in photoproduction of a heavy meson or a $Z$, one can have cuts of $H$ in both $(q+k)^2$ and $(q-k)^2$ at the same value of $x$ and needs both representations of $M$ at the same time.

In the case of meson production (Fig. 1(b)) one can also cut through the blob representing the meson distribution function, with one valence quark at either side of the cut, or one can cut through one of the valence quark lines. In this case one will use that the meson distribution amplitude can be written in terms of singularities in one of the quark virtualities $l^2$ and $l'^2$ as discussed in Section 7.

**Acknowledgements**

We gratefully acknowledge discussion with Jochen Bartels, Martin Beneke, John Collins, Peter Landshoff, Piet Mulders, Otto Nachtmann and Bernard Pire. Special thanks are due to O. Nachtmann for reading the manuscript. M.D. would like to thank the NIKHEF Theory Group for its hospitality.

T.G. was carrying out his work as part of a training Project of the European Community under Contract No. ERBFMBI-CT95-0411. This work has been partially funded through the European TMR Contract No. FMRX-CT96-0008: Hadronic Physics with High Energy Electromagnetic Probes. CPT is Unité Propre du CNRS and SUBATECH is Unité Mixte de l’Université de Nantes, de l’Ecole des Mines de Nantes et du CNRS.

**Appendix A**

In this appendix we study in some detail the identity of the matrix elements $M$ and $\hat{M}$ introduced respectively in (10) and (15). As we already mentioned the $k_x$ integration has no relevance in the derivation so that we perform our reasoning on the corresponding $k_{\perp}$ unintegrated quantity. We focus on the most complicated region $0 < x < \xi$ where as discussed in point C. of Section 5 the integration over $k^-$ inevitably picks up singularities in more than one invariant and where the cutting rules have to be applied with some care. In order to deal with complete but readable expressions we ignore here complexities due to the spin of both partons and hadrons and introduce

$$\mathcal{M} = \int_{-\infty}^{+\infty} dk^- \int d^4 z e^{ik\cdot z} \langle p' | T \phi^+(0) \phi(\zeta) | p \rangle, \quad \hat{\mathcal{M}} = \int_{-\infty}^{+\infty} dk^- \int d^4 z e^{ik\cdot z} \langle p' | \phi^+(0) \phi(\zeta) | p \rangle,$$

(A.1)

where $\phi$ is a charged scalar ‘‘quark’’ field and $|p\rangle$ a scalar ‘‘proton’’ state.

In the framework we are using we invoke perturbation theory for the purpose of analytic properties and of applying the cutting rules. To be consistent we must treat quarks as free particles with a mass $m_q$. The scalar analogue of the amplitude $\mathcal{M}$ in (11) then has cuts and a mass pole in each of $k^2$ and $k'^2$. To apply the cutting rules we will isolate the poles from the rest of the singularity structure, and also keep track of possible disconnected terms of the matrix elements that appear in our argument. To this end we use the framework of the reduction formula [13] (cf. also [14]) and introduce an interpolating field $\Phi$ for the proton. For $p \neq p'$ the matrix

---

12 Note that such cuts already appear if one takes the hard blob $H$ at Born level, cf. e.g. the diagrams in Fig. 2 of [4].
element $\langle p'| T\phi^\dagger(0) \phi(z) | p \rangle$ does not have a disconnected part, and in the case of diagonal parton densities one explicitly only takes its connected part. One thus has

$$\int d^4 z e^{i k \cdot z} \langle p'| T\phi^\dagger(0) \phi(z) | p \rangle = \frac{i \sqrt{Z_q}}{k^2 - m_q^2 + i \varepsilon} \frac{i \sqrt{Z_p}}{k^2 - m_p^2 + i \varepsilon} i \tilde{G}( p \bar{q} \to p' \bar{q}') ,$$  

(A.2)

with

$$(2\pi)^4 \delta^{(4)}( p - k - p' + k') \cdot i \tilde{G} = \frac{1}{Z_q} \frac{1}{Z_p} \int d^4 z d^4 z' d^4 y d^4 y' e^{i k \cdot z - i k' \cdot z'} e^{i p \cdot y - i p' \cdot y'} \cdot \left( \frac{\mathcal{D} + m_q^2}{\mathcal{D} + m_q^2} \right) \left( \frac{\mathcal{D} + m_p^2}{\mathcal{D} + m_p^2} \right) \left( \mathcal{D} + m_p^2 \right)$$

$$(2\pi)^4 \delta^{(4)}( p - k - p' + k') \cdot i \tilde{G} = \frac{1}{Z_q} \frac{1}{Z_p} \int d^4 z d^4 z' d^4 y d^4 y' e^{i k \cdot z - i k' \cdot z'} e^{i p \cdot y - i p' \cdot y'} \cdot \langle 0 | T\phi^\dagger( z') \phi( z) \Phi( y') \Phi( y) | 0 \rangle ,$$  

(A.3)

where $Z_q$ and $Z_p$ are the wave function normalisation constants of the quark and the proton, and where four-momenta are labelled as in (12). In the limit where both parton legs go on shell $\tilde{G}$ is a $\mathcal{F}$-matrix element: depending on the signs of $k^0$ and $k'^0$ the partons are in the initial or final state. We emphasise that only the parton mass poles in $k^2$ or $k'^2$ have been isolated from $\tilde{G}$ which thus still contains the branch cuts in these variables, unlike truncated Green functions where the full propagators including these cuts are split off.

To obtain $\mathcal{M}$ we integrate the right hand side of Eq. (A.2) over $k^-$ and close the contour in the upper half plane, assuming that convergence is fast enough at infinity. In the region $0 < x < \xi$ the only singularities in the upper half $k^-$ plane are due to the pole $1/(k^2 - m_q^2 + i \varepsilon)$ and the cuts in $s$ and $k'^2$ of $\tilde{G}$, according to our hypothesis on the singularity structure. We get

$$\mathcal{M} = \int_{-\infty}^{+\infty} dk^- 2\pi \delta(k^2 - m_q^2) \sqrt{Z_q} \frac{i \sqrt{Z_q}}{k^2 - m_q^2 + i \varepsilon} i \tilde{G}( p \bar{q} \to p' \bar{q}')$$

$$+ \int_{-\infty}^{+\infty} dk^- \frac{i \sqrt{Z_q}}{k^2 - m_q^2 - i \varepsilon} \frac{i \sqrt{Z_q}}{k^2 - m_q^2 + i \varepsilon} i \left( \text{disc}_x \tilde{G} + \text{disc}_x^{\star} \tilde{G} \right) .$$  

(A.4)

Note that the sign of $i \varepsilon$ in the $k^2$ pole has changed from (A.2) to (A.4) as a consequence of separating the contributions of the pole and the cuts, with the pole lying below the cuts in the $k^-$ plane.

Using the cutting rules [15] (cf. also [14]) the above discontinuities can be expressed as

$$- i \text{disc}_x \tilde{G}( p \bar{q} \to p' \bar{q}') = \sum_X (2\pi)^4 \delta^{(4)}( p_X + k' - p') \tilde{G}( p \bar{q} \to X) \tilde{G}^\star( p' \bar{q}' \to X) ,$$

(A.5)

$$- i \text{disc}_x^{\star} \tilde{G}( p \bar{q} \to p' \bar{q}') = \sum_{X' \neq X'} (2\pi)^4 \delta^{(4)}( p_X + k' - p') \tilde{G}( p \bar{q} \to p' X') \tilde{G}^\star( \bar{q}' \to X') ,$$

(A.6)

where $p_X$ and $p_X'$ respectively denote the four-momenta of the cut states $X$ and $X'$. The functions $\tilde{G}$ are defined in analogy to (A.3), i.e. as Fourier transformed Green functions with the poles in the external legs removed. They reduce again to $\mathcal{F}$-matrix elements in the on-shell limit of the external antiquark legs. The $k^2$ cut (A.6) has no contribution from a single antiquark state, as expressed in the restriction $X' \neq \bar{q}'$, where $\bar{q}'$ has momentum $-k^2$: if $\bar{q}'$ is off-shell such a state is excluded by momentum conservation and if it is on-shell then the amplitude $\tilde{G}$ is zero for the non-interacting transition $\bar{q}' \to \bar{q}'$. 


Let us now turn to the study of $\hat{M}$ when $0 < x < \xi$. Inserting a complete set of “out” states\(^\text{13}\) we obtain

$$\hat{M} = \int_{-\infty}^{+\infty} dk^- \sum_X (2\pi)^4 \delta^{(4)}( p_X + k^- - p') \frac{i\sqrt{Z_q}}{k^2 - m^2 + i\varepsilon} i\tilde{G}( p\bar{q} \rightarrow X).$$

Due to momentum conservation $|X\rangle$ cannot contain the initial proton state $|p\rangle$ in the region $0 < x < \xi$ we are considering. The second matrix element in (A.7) is then connected and can be written in terms of $\tilde{G}$ using again the reduction formula:

$$\text{out} \langle X|\phi(0)|p\rangle_{\text{in}} = \frac{i\sqrt{Z_q}}{k^2 - m^2 + i\varepsilon} i\tilde{G}( p\bar{q} \rightarrow X).$$

There are however states that contain the final proton state $|p'\rangle$ and lead to disconnected parts of

$$\text{out} \langle p'X|\phi(0)|p\rangle_{\text{in}} = \frac{i\sqrt{Z_q}}{k^2 - m^2 + i\varepsilon} i\tilde{G}( p\bar{q} \rightarrow p'X') \left(2\pi \right)^3 2 p^0 \delta^{(3)}(p' - p') \frac{i\sqrt{Z_q}}{k^2 - m^2 + i\varepsilon} i\tilde{G}(\bar{q} \rightarrow X').$$

If $X'$ is a single antiquark state $\bar{q}'$, which due to momentum conservation requires $\bar{q}'$ to be on shell, the disconnected term at the r.h.s. of (A.9) simply reads

$$(2\pi)^3 2 p^0 \delta^{(3)}(p' - p') \langle X' | \phi(0) | p\rangle_{\text{in}} = (2\pi)^3 2 p^0 \delta^{(3)}(p' - p') \frac{i\sqrt{Z_q}}{k^2 - m^2 + i\varepsilon} i\tilde{G}( p\bar{q} \rightarrow p'X').$$

For states $|X\rangle$ that do not contain a proton $|p'\rangle$ the analogue of (A.8) holds. Together with the connected term in (A.9) these states give the $s$-channel cut in $\hat{M}$, whereas the disconnected terms give the cuts and pole in $k^2$. Using

$$\sum_{X' = \bar{q}'\ldots} (2\pi)^4 \delta^{(4)}( p_{X'} + k') = 2\pi \delta(k^2 - m^2)$$

for the pole term we obtain

$$\hat{M} = \int_{-\infty}^{+\infty} dk^- \left[2\pi \delta(k^2 - m^2) \frac{i\sqrt{Z_q}}{k^2 - m^2 + i\varepsilon} i\tilde{G}( p\bar{q} \rightarrow p'\bar{q}') \right. \left. + \sum_{X' = \bar{q}'\ldots} (2\pi)^4 \delta^{(4)}( p_{X'} + k') \frac{i\sqrt{Z_q}}{k^2 - m^2 + i\varepsilon} \frac{-i\sqrt{Z_q}}{k^2 - m^2 - i\varepsilon} i\tilde{G}( p\bar{q} \rightarrow p'X') \tilde{G}^\dagger(\bar{q} \rightarrow X') \right] \left. + \sum_X (2\pi)^4 \delta^{(4)}( p_X + k^- - p') \frac{i\sqrt{Z_q}}{k^2 - m^2 + i\varepsilon} \frac{-i\sqrt{Z_q}}{k^2 - m^2 - i\varepsilon} i\tilde{G}( p\bar{q} \rightarrow X) \tilde{G}^\dagger( p'\bar{q}' \rightarrow X) \right).$$

With (A.4) and the cutting rules (A.5), (A.6) we finally have $\hat{M} = \hat{M}$, and thus $\hat{M} = M$.\(^{14}\)

---

\(^{13}\) So far we only had matrix elements between one-particle states or the vacuum, which are equal in the “in” and “out” representations, here we need to make the distinction.

\(^{14}\) In the region $\xi < s < 1$ discussed in point A. of Section 5 some attention is required to see that the signs of $i\varepsilon$ in the $k^2$ and $k'^2$ poles are compatible between $\hat{M}$ and $\hat{M}$. Assuming that disc $\tilde{G}$ vanishes for $s$ below $m^2$ due to its threshold one finds however that these signs can actually be chosen arbitrarily.
References


NLO QCD corrections and triple gauge boson vertices at the NLC

K.J. Abraham a, Bodo Lampe b

a Department of Physics, University of Natal, Pietermaritzburg, South Africa
b Max Planck Institut für Physik, Föhringer Ring 6, D-80805 München, Germany
and Department of Physics, University of Munich, Theresienstr. 37, D-80833 München, Germany

Received 3 February 1998; revised 26 March 1998
Editor: R. Gatto

Abstract

We study NLO QCD corrections as relevant to hadronic $W$ decay in $W$ pair production at a future 500 GeV $e^+e^-$ linac, with particular emphasis on the determination of triple gauge boson vertices. We find that hard gluon bremsstrahlung may mimic signatures of anomalous triple gauge boson vertices in certain distributions. The size of these effects can strongly depend on the polarisation of the initial $e^+e^-$ beams. © 1998 Elsevier Science B.V. All rights reserved.

Although the Standard Model is in excellent agreement with existing collider data, there are strong grounds to expect discrepancies to appear once future high energies accelerators are commissioned. In particular, Physics Beyond the Standard Model may show up in the form of anomalous triple gauge boson vertices. Although LEP data will constrain these trilinear couplings, high statistics at energies far above threshold are needed to pinpoint small deviations from the Standard Model. Thus detailed analyses have been performed on the sensitivity of $W$ pair production and decay at future high energy $e^+e^-$ linacs [1−3] to the presence of anomalous triple gauge boson vertices. These studies, however, have not taken into account NLO QCD effects in hadronic $W$ decay (see however [4]). In this letter, we will argue that these effects can generate deviations from tree level Standard Model predictions which are large enough to influence discovery bounds on anomalous form factors.

The differential cross-section for $W$ pair production and decay (in the narrow width approximation which we use throughout following [5,1]) can be schematically written as

$$d\sigma \sim \frac{\Gamma_a \Gamma_b}{\Gamma \tau} \sum_{\lambda \tau_1 \tau_2 \tau_3 \tau'_3} F_{\tau_3 \tau_3}^\lambda (s, \cos \theta) F_{\tau_1 \tau_1}^{\lambda'}(s, \cos \theta) \times D_{\tau_1 \tau_2} D_{\tau_1 \tau_2'}$$

where $F$ is a generic helicity amplitude dependent on $\sqrt{s}$ and production angle $\theta$, and $D_{\lambda\tau}$ denotes a generic element of the density matrix for $W$ decay. $\lambda$ denotes the electron helicity (± ½) while $\tau$ denotes $W$ helicities (+, −0). $\Gamma$ is the total width of the $W$ while $\Gamma_a$ and $\Gamma_b$ denote partial widths to the final states of interest. The precise forms of $F$ and $D$ may be found in [5]. The constants we have not explicitly written in Eq. (1) are purely kinematical overall factors common to all final states and which

PII: S0370-2693(98)00408-0

0370-2693/98/$19.00 © 1998 Elsevier Science B.V. All rights reserved.
are not relevant for the numerical results we will present later on.

For the sake of convenience, we reproduce the diagonal elements of $D$ from [5]

$$D_{++} = \frac{1}{2}(1 + \cos^2 \theta) - \{\cos \theta\}$$

(2)

$$D_{--} = \frac{1}{2}(1 + \cos^2 \theta) + \{\cos \theta\}$$

(3)

$$D_{00} = \sin^2 \theta$$

(4)

where $\theta$ is the polar angle of the outgoing fermion in the rest frame of the decaying $W$ and all fermions are assumed massless. We will assume that the rest frame of each $W$ can be reconstructed, thereby excluding purely leptonic final states. For hadronic $W$ decays, where the jet charges cannot be reconstructed, symmetrisation between quarks and antiquarks requires that the terms within $\{\}$ must be dropped. The off diagonal elements of $D$ depend, in addition to $\theta$, on the azimuthal angle $\phi$, but make no contribution to the total cross-section. They are however relevant for azimuthal correlations, which we will not discuss.

Once NLO QCD effects are included in $W$ decays, the formulae above must be modified. In addition to $\theta$ and $\phi$, the matrix elements for gluon bremsstrahlung contributions depend on additional phase space variables, and are singular in the collinear limit, quite apart from divergences due to virtual corrections. As these singularities cancel only when IR safe quantities are calculated, IR safe generalisations of $\theta$ and $\phi$ are required. One way of proceeding follows from the observation that $\theta$ in Eq. (4), defined in terms of the quark direction, is at LO also the polar angle of the thrust direction. As thrust is IR safe [6], NLO $D$ functions defined in terms of the thrust orientation, which for our purposes is the direction of the most energetic outgoing parton in the $W$ rest frame, are guaranteed to be singularity free.

Restricting ourselves once again to diagonal matrix elements we have at NLO [7]

$$D_{AA} = \left(1 + \frac{\alpha_s}{\pi}\right)\left(1 - 3 LC_F \frac{\alpha_s}{2\pi}\right)$$

$$\times \left(D_{00} + 2 LC_F \frac{\alpha_s}{2\pi}\right)$$

(5)

neglecting terms of $\sigma(\alpha_s^2)$. $D^0$ denotes the symmetrised leading order term and $L$ is a numerical constant of value 0.4875 which is a relic of numerical integration over Dalitz variables. All angles in Eq. (5) now refer not to a given outgoing parton but to the thrust axis. As before, all outgoing partons are assumed massless. It should be noted that there are additional terms linear in $\cos \theta$ in Eq. (5) which vanish once we assume that jet charges are not determined, and which we have therefore dropped.

Retaining terms to $\sigma(\alpha_s)$ only

$$\sum_A D_{AA} = \left(1 + \frac{\alpha_s}{\pi}\right)\sum_A D^0_{AA}$$

The term in brackets is the well known $\sigma(\alpha_s)$ QCD K factor for hadronic $W$ decay, as expected. We thus see, that for observables for which polarisation is not relevant, the NLO corrections may be obtained by rescaling LO results by a constant factor which we will derive shortly. In the Standard Model where gauge cancellations ensure the suppression of longitudinal gauge boson production at high energies, polarisation may be relevant for some observables, for example asymmetries, where different polarisation states are in general weighted differently. Hence we will not only approximate NLO effects by a constant K factor, but will also make use of Eq. (5) convoluted with Eq. (1).

Before proceeding further along these lines, it is useful to study the possible significance of NLO QCD effects in the analysis of anomalous gauge boson vertices. To do so, it is instructive to consider the term within $\{\}$ in Eq. (5) for $A = +$. This term can be rewritten as

$$\frac{(1 + \cos^2 \theta)}{2} \left(1 + 2 LC_F \frac{\alpha_s}{2\pi}\right) + \frac{\sin^2 \theta}{2} - 2 LC_F \frac{\alpha_s}{2\pi}$$

For transversely polarised gauge bosons, the LO distribution gets rescaled and in addition a longitudinally polarised component seems to appear. However, the appearance of additional longitudinal modes is one of the hallmarks of non-standard triple gauge boson vertices! Thus we see that NLO QCD has the potential to mimic signatures of anomalous triple gauge boson vertices. The size of this effect is proportional to $L$, indicating that hard gluon bremsstrahlung is responsible. It is significant that in Eq. (5) $\alpha_s$ is evaluated at $M_W$ independent of $\sqrt{s}$.


NLO QCD corrections thus do not diminish in size with increasing energy, in contrast to other radiative corrections such as initial state bremsstrahlung and finite width contributions.

For longitudinal modes \((A = 0)\) it is easy to see from Eq. (4) and Eq. (5) that the QCD \(K\) factor at \(\theta = 0\) is infinite. Thus large QCD corrections may be expected in distributions where longitudinal modes are important. The infinite \(K\) factor is due to the vanishing of the LO cross-section at \(\theta = 0\), which may be understood in terms of angular momentum conservation, and has been observed in other processes involving hadronic \(W\) decay [8].

To derive expected \(K\) factors, it is important to note that NLO QCD effects also appear in the redefinition of the width and branching fractions of the \(W\), which appear in Eq. (1). As we will focus on final states containing leptons and hadrons, what appears at Born level is (in obvious notation), \(\Gamma_L/\Gamma_H/\Gamma^2\) where all widths are calculated from tree level expressions. Keeping terms up to and including \(\mathcal{O}(\alpha_s)\) only \(^1\)

\[
\Gamma^2 \to \Gamma^2 \left(1 + 2\frac{2}{3} \frac{\alpha_s}{\pi}\right)
\]

The NLO cross-section with no phase space cuts can be obtained from Born level by making a further change i.e.

\[
\Gamma_H \to \left(1 + \frac{\alpha_s}{\pi}\right)\Gamma_H
\]

Thus the change to the cross-section can be accounted for by rescaling by a factor \(\mathcal{R}\) given by

\[
\mathcal{R} = \left(1 - \frac{1}{3} \frac{\alpha_s}{\pi}\right)
\]

neglecting terms with higher powers of \(\alpha_s\). This is the constant factor mentioned earlier.

As we have argued earlier, the sum total of NLO QCD corrections may not always be obtained by rescaling by the constant factor in Eq. (6). Therefore we will define a \(C\), a modified \(K\) factor as follows;

\[
C = \frac{\text{NLO} - \mathcal{R}\text{LO}}{\mathcal{R}\text{LO}}
\]

If \(C\) vanishes or is very small for a certain observable in a given region of phase space, then a rescaling is sufficient to describe NLO effects. If this is not the case, then NLO effects are significant. Similarly, we can define \(C_A\) to describe the corrections due to anomalous couplings as follows,

\[
C_A = \frac{(A - \text{LO})}{\text{LO}}
\]

where \(A\) is the contribution for a certain choice of anomalous triple gauge boson vertices.

Table 1

<table>
<thead>
<tr>
<th>(\cos \theta)</th>
<th>(\theta)</th>
<th>(\theta)</th>
<th>(\theta)</th>
<th>(\theta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>.8</td>
<td>-.012</td>
<td>-.010</td>
<td>-.002</td>
<td>.006</td>
</tr>
<tr>
<td>.000</td>
<td>.007</td>
<td>.027</td>
<td>.050</td>
<td></td>
</tr>
<tr>
<td>-.4</td>
<td>-.016</td>
<td>-.013</td>
<td>-.002</td>
<td>.013</td>
</tr>
<tr>
<td>.000</td>
<td>.002</td>
<td>.007</td>
<td>.016</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>(\infty)</td>
<td>.081</td>
<td>-.001</td>
<td>-.014</td>
</tr>
<tr>
<td>-.061</td>
<td>-.061</td>
<td>-.061</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.4</td>
<td>-.013</td>
<td>-.010</td>
<td>-.002</td>
<td>.007</td>
</tr>
<tr>
<td>.000</td>
<td>-.001</td>
<td>-.005</td>
<td>-.009</td>
<td></td>
</tr>
<tr>
<td>.8</td>
<td>-.014</td>
<td>-.011</td>
<td>-.002</td>
<td>.009</td>
</tr>
<tr>
<td>.000</td>
<td>.000</td>
<td>-.001</td>
<td>-.002</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 2. The values of $C$ are plotted as a function of $\cos \theta$ and $\theta_y$ for the variable defined in Eq. (9). The resulting surface represents the QCD part of Table 1, i.e. the upper figure in each entry of Table 1.

We will now present numerical results for various observables at LO and NLO to illustrate the size of the corrections to be expected. We fix throughout, $\sqrt{s} = 500$ GeV, $M_Z = 91.187$ GeV, $M_W = 80.33$ GeV, $\alpha = 1/128$, $\sin \theta_W = .23$, and $\alpha_s = .12$. For the sake of definiteness, we assume that the $W^-$ decays hadronically, and the $W^+$ leptonically. The incoming $e^+e^-$ beams are unpolarised unless otherwise stated, and the anomalous contributions are defined via Eq. (8).

Angles referring to outgoing fermions and jets are defined in the rest frame of the parent $W^\pm$.

To begin with, we demonstrate the utility of absorbing $K$ into the LO cross-section by studying the $C$ dependence of the differential cross-section with respect to $\theta_y$, the polar angle of the thrust axis. The two curves in Fig. 1 correspond to $C$ with $K$, defined by Eq. (6) and $K = 1$. We see immediately that in this instance a significant part of the NLO corrections can be absorbed into the redefinition of widths, thereby reducing the magnitude of such effects.

However, this is not always the case. For example, for observables for which the LO contributions vanish in certain regions of phase space, NLO effects cannot be simply accounted for by a redefinition of widths. One such observable discussed in [7] is given by

$$\int_{-1}^{1} d\cos \theta_+ \frac{\cos \theta_+}{|\cos \theta_+|} \frac{d\sigma (e^+ e^- \rightarrow f^+ \nu f^- X)}{d\cos \theta_d d\cos \theta_\perp d\cos \theta_\parallel}$$

which corresponds to the double differential distribution with respect to $\cos \theta$ and $\cos \theta_\perp$, with the azimuthal angles integrated over and the polar angle of the charged anti-lepton integrated over anti-symmetrically.

This observable may seem rather contrived, however, being asymmetric by construction it is sensitive to the C and P Violating form factor denoted by $z_\perp$ in [5]. As can be seen from Table 1 and Figs. 2 and 3, NLO QCD effects and a non zero value of $z_\perp$ both generate appreciable corrections to the Standard Model predictions for the distribution described in Eq. (9), particularly for $\cos \theta \sim 0$. The need for taking into account NLO QCD effects in the analysis of anomalous triple gauge boson vertices is apparent.

Table 2
The values of $C$ and $C_\perp$ are plotted as a function of $\cos \theta$ and $\theta_y$ for the double differential cross-section evaluated at $\theta_y = 1$ with polarised beams ($e^+\gamma$ and $e^-\gamma$). The upper figure in each entry is $C$, while the lower figure is $C_\perp$ evaluated for anomalous parameters described in the text. $\cos \theta$ runs along the vertical axis and $\theta_y$ (in units of $\pi$) along the horizontal axis. All azimuthal angles have been integrated over.

<table>
<thead>
<tr>
<th>$\cos \theta$</th>
<th>0</th>
<th>.15</th>
<th>.30</th>
<th>.45</th>
</tr>
</thead>
<tbody>
<tr>
<td>.8</td>
<td>1.551</td>
<td>.218</td>
<td>.085</td>
<td>.058</td>
</tr>
<tr>
<td>.9</td>
<td>.290</td>
<td>.216</td>
<td>.209</td>
<td>.207</td>
</tr>
<tr>
<td>-.4</td>
<td>.753</td>
<td>.269</td>
<td>.191</td>
<td>.173</td>
</tr>
<tr>
<td>0</td>
<td>.360</td>
<td>.226</td>
<td>.205</td>
<td>.201</td>
</tr>
<tr>
<td>.4</td>
<td>.386</td>
<td>.256</td>
<td>.233</td>
<td>.229</td>
</tr>
<tr>
<td>.393</td>
<td>.232</td>
<td>.203</td>
<td>.198</td>
<td></td>
</tr>
<tr>
<td>.4</td>
<td>.175</td>
<td>.165</td>
<td>.166</td>
<td>.169</td>
</tr>
<tr>
<td>.412</td>
<td>.227</td>
<td>.205</td>
<td>.201</td>
<td></td>
</tr>
<tr>
<td>.8</td>
<td>.038</td>
<td>.043</td>
<td>.045</td>
<td>.051</td>
</tr>
<tr>
<td>.424</td>
<td>.216</td>
<td>.209</td>
<td>.207</td>
<td></td>
</tr>
</tbody>
</table>

Note that we have included both a tabular (Table 1) and a graphical (Figs. 2 and 3) representation of our results. The former allows to read off easily precise numbers, whereas the latter is good for a global overview of the results.

As has been pointed out in [5] and [1], triple differential distributions are particularly sensitive to anomalous gauge boson couplings. Hence, as a further illustration of the relevance of Eq. (5) we now sample a triple differential distribution with all azimuthal angles integrated over and the polar angle of the thrust axis is fixed at 0.1. In addition, the incoming beams are polarised \((e^+_R, e^-_L)\). The non-zero anomalous couplings are chosen to be (in the notation of [5])

\[
x_y = 0.005, \quad \delta_\gamma = \frac{x_y}{\sin \theta_W \cos \theta_W}, \quad x_\gamma = -x_y \frac{\sin \theta_W}{\cos \theta_W}
\]

This choice of parameters is motivated by the scenario in [9] and the values above are slightly above the threshold for discovery at \(\sqrt{s} = 500\) GeV according to LO analyses in [1].

Our results for these choices are given in Table 2 and Figs. 4 and 5. They clearly indicate that although the anomalous couplings produce sizable deviations from the tree level predictions of the Standard Model, NLO QCD effects are definitely not negligible in comparison and need to be taken into account to establish discovery limits. Note that we have included both a tabular (Table 2) and a graphical (Figs. 4 and 5) representation of our results. The former allows to read off easily precise numbers, whereas the latter is good for a global overview of the results.

It is worth noting that for opposite incoming beam helicities, both the anomalous corrections and NLO QCD corrections in the same region of phase space are much smaller; the NLO QCD corrections are never more than a few percent. This can be understood from the fact that for incoming \(e_y^+\), the outgoing \(W^+\) is largely longitudinally polarised, while for incoming \(e_y^-\) the outgoing \(W^-\) is largely transverse. This strong dependence of the size of NLO QCD effects on the incoming beam polarisation has not been pointed out before, and is particularly significant, as several authors have suggested beam polarisation as a diagnostic tool to unravel the structure of anomalous gauge boson interactions [2,10]. It is also noteworthy that Table 1 and Figs. 2 and 3 are very different from Table 2 and Figs. 4 and 5, indicating that NLO QCD corrections to different observables may not be simply obtained from some universal prescription, but must be calculated from scratch.

Common to the results presented in Tables 1 and 2, apart from the strong dependence of \(K\) factors with phase space, is the fact that different polarisation states make different contributions to the distributions under consideration, either due to a choice of initial polarisation or due to an observable being asymmetric by construction. As we have argued earlier it is precisely in such cases that NLO QCD

---

**Fig. 4.** The values of \(C\) are plotted as a function of \(\cos \theta\) and \(\theta_A\) for the double differential cross-section evaluated at \(\theta \rightarrow 1\) with polarised beams \((e_R^+\) and \(e_L^-\)). The resulting surface represents the QCD part of Table 2, i.e. the upper figure in each entry of Table 2.

**Fig. 5.** The values of \(C_\gamma\) are plotted as a function of \(\cos \theta\) and \(\theta_A\) for the double differential cross-section evaluated at \(\theta \rightarrow 1\) with polarised beams \((e_R^+\) and \(e_L^-\)). The resulting surface represents the anomalous part of Table 2, i.e. the lower figure in each entry of Table 2.
corrections could be non-trivial and this is indeed consistent with our results, and with Figs. 1–3 of [7] where a sizable variation in \(K\) factors is also observed. This suggests a useful rule of thumb; higher order QCD corrections should not be approximated by constant \(K\) factors where \(W\) polarisation is observed and/or where different polarisation states make different contributions to the observables under consideration. Finally, it is worth repeating that the magnitude of the relevant \(K\) factors is controlled by \(\alpha_s(M_W)\), and is thus independent of \(\sqrt{s}\) insofar as the narrow width approximation is valid. It is not surprising therefore, that the broad features of Tables 1 and 2 and Figs. 2–5 persist at higher energies as well.

To summarise, we have demonstrated the importance of NLO QCD corrections in the analysis of triple gauge boson vertices at future \(e^+e^-\) linacs. The magnitude of these corrections varies strongly with beam polarisation and seem to be particularly large for asymmetries, and certainly will affect the exclusion bounds for anomalous triple gauge boson vertices. A precise quantitative estimate will be possible only with a detailed analysis including detector acceptances which is beyond the scope of this letter, but is definitely worth undertaking.

K.J.A. wishes to thank Jose Wudka for valuable clarifications and encouragement. B.L. acknowledges useful discussions with J.G. Körner.

References

Probing the BFKL gluons with $J/\psi$ leptoproduction

S.P. Baranov 1
P.N. Lebedev Institute of Physics, Leninsky prosp. 53, Moscow 117924, Russia

Received 2 December 1997; revised 5 March 1998
Editor: P.V. Landshoff

Abstract

We analyse the production of $J/\psi$ particles in high energy electron proton collisions paying attention to off shell gluon states in the photon-gluon fusion subprocess. We find that the degree of $J/\psi$ spin alignment is sensitive to the initial gluon virtuality. The considered process may be proposed as a test of noncollinear parton evolution theory.

q 1998 Published by Elsevier Science B.V. All rights reserved.

Keywords: Semihard QCD; Quarkonia production

1. Introduction

At the energies of modern colliders, the interaction dynamics is governed by the properties of parton distributions in the small $x$ region. This domain is characterized by the double inequality

$s \gg \mu^2 = \sqrt{s} \gg \Lambda^2$,

which shows that the typical parton interaction scale $\mu$ is much higher than the QCD parameter $\Lambda$ but is much lower than the total c.m.s. energy $\sqrt{s}$. The situation is therefore classified as "semihard". In such a case, the perturbative QCD expansions in $\alpha_s$ may contain large coefficients $O(\ln(s/\mu^2)) = O(\ln(1/x))$ which compensate the smallness of the coupling constant $\alpha_s(\mu^2/\Lambda^2)$. The resummation [1-4] of the terms $[\ln(\mu^2/\Lambda^2) \alpha_s]$, $[\ln(\mu^2/\Lambda^2) \ln(1/x) \alpha_s]$ and $[\ln(1/x) \alpha_s]^2$ results in the parton distributions $\Phi_i(x,k_t^2,\mu^2)$ that generalize the factorization of the hadronic matrix elements beyond the collinear approximation (hereafter these generalized distribution functions will be referred to as "noncollinear"). The distributions $\Phi_i(x,k_t^2,\mu^2)$ determine the probability to find a parton of type $i$ carrying the longitudinal momentum fraction $x$ and transverse momentum $k_t$ at the probing scale $\mu^2$. They obey the BFKL [5] (Balitsky-Fadin-Kuraev-Lipatov) equation and reduce to the conventional parton densities once the $k_t$ dependence is integrated out:

$$\int_0^{\mu^2} \Phi_i(x,k_t^2,\mu^2) d k_t^2 = x F_i(x,\mu^2).$$

Nowadays, the significance of noncollinear approach becomes more and more commonly recognized. Its applications to a variety of photo-, lepto- and hadroproduction processes may be found in [6-8]. To search for a direct experimental evidence for the validity of semihard theory is an intriguing and topical problem.

---

1 Electronic address: baranov@sgi.lpi.msk.su.
The specific properties of semihard theory may manifest in several ways. First, it is the non strong $k_T$ ordering in the parton evolution, contrasting to the strong $k_T$ ordering implied in the collinear factorization approach [9]. An experimental observation of this effect would require a jet-jet correlation analysis. With respect to inclusive production properties, one points out an additional contribution to the cross sections due to the integration over the $k_T^2$ region above $\mu^2$ and the broadening of the $p_T$ spectra due to extra transverse momentum of the colliding partons. It is the matter of principal importance that the interacting partons are not on mass shell but are characterized by virtual masses proportional to their transverse momenta: $m^2 = -k_T^2/(1-x)$. This also assumes a modification of the polarization density matrix. In particular, the polarization vector of a gluon is no longer purely transversal, but acquires an admixture of longitudinal and time-like components.

In the present note we analyse the production of $J/\psi$ particles in high energy electron-proton collisions. We find that the degree of $J/\psi$ spin alignment is sensitive to the initial gluon off-shellness. Thus, the considered process may be useful as a test of BFKL evolution principles.

2. Model

In the framework of colour singlet model [10], the production of a heavy vector meson is described as the perturbative production of a colour singlet $q\bar{q}$ pair with the quantum numbers of the quarkonium state. The formation of a meson from the quark pair is a non-perturbative process. Its probability reduces to a single parameter related to the meson wave function at the origin $|\Psi(0)|^2$, which is known for $J/\psi$ and $\Upsilon$ families from the measured leptonic decay widths $\Gamma_{\ell\ell}$. This model yields a satisfactory description [11] of recent high energy experimental data [12,13]. Here we notice the theoretical advantage of $J/\psi$ leptoproduction in comparison with hadroproduction: the latter seems to be dominated by other, large and poorly calculable contributions [14–16].

As the available experimental data are well described by the colour singlet model alone, it may look unreasonable to include other production mechanisms, such as the colour octet model, although the latter is widely discussed in the literature. We also can point out that the colour octet model is not free from internal difficulties. The alternative model [15] has been shown to bring only a negligible contribution to $J/\psi$ leptoproduction. Playing for safety, we, nevertheless, will restrict to certain kinematical cuts to avoid ambiguous regions.

In the present paper we extend the colour singlet model to off-shell gluon states. Fig. 1 displays one of the six Feynman diagrams corresponding to the leading order QCD matrix element, where the virtual gluon polarization matrix takes the form [4]:

$$\varepsilon^{\mu\nu} = k_T^\mu k_T^\nu/(k_T^2)^2.$$  \hspace{1cm} (3)

The above prescription has a clear analogy with the well known equivalent photon approximation. Consider for example a photon emitted by an electron: $e(p) \rightarrow e'(p') + \gamma(k)$. Then, taking trace over the electron line in the matrix element squared one obtains the polarization tensor $L^{\mu\nu}$:

$$L^{\mu\nu} = \text{tr}\left\{(\not{\omega} + m_e)(\not{\omega} + m_e)\gamma^\mu (\not{\omega} + m_e)\gamma^\nu\right\}$$

$$= 8 \not{p}\not{p} - 4(p\cdot k)g^{\mu\nu}. \hspace{1cm} (4)$$

Neglecting the second term in the right hand side in the small $x$ limit, $p \gg k$, one immediately arrives at the spin structure $\varepsilon^{\mu\nu} \sim L^{\mu\nu} \sim (p\cdot p)/(p\cdot k)$ and to apply a gauge shift $\epsilon^\mu \rightarrow \epsilon^\mu - k^\mu/x$.

Another essential model ingredient is the non-collinear gluon distributon. The question of parametrizing the parton distributions $\Phi(x, k^2, \mu^2)$ remains yet open because the appropriate analysis of the experimental data has never been carried out. For the present purposes, we find it reasonable to follow

---

Footnote: With the colour octet matrix elements [17] fitted to hadroproduction data, the colour octet model predictions are at odds [18] with electroproduction data because of extremely strong peak at large $z = E_\gamma/E_e$. The contradiction may be softened [16] if to take into account the initial parton transverse motion. Then one needs much lower matrix element values to fit the hadroproduction, thus also greatly reducing the colour octet contribution to electroproduction. On the other hand, the calculations in [16] are not done in a consistent way, as they include $O(a_s^3)$ terms and ignore $O(a_s^2)$ terms. Being added, the latter would enhance the cross section at low $p_T$, that, in turn, conflicts with hadroproduction data.
the prescriptions of paper [19]. The proposed method lies upon a straightforward perturbative solution of the BFKL equation where the collinear gluon density \( x G(x, \mu^2) \) is used as the boundary condition. This automatically fulfills the condition (2), guarantees the validity of BFKL evolution, and conveniently provides a common smooth parametrization for the whole range of \( k_T \) without any extra parameters taken from outside. The described method has been implemented in our simulations with the GRV structure functions set [20] taken for the input collinear densities. We have not included any screening corrections discussed in the literature [21] since they are expected to be unimportant up to HERA energies.

3. Details of calculations

Let \( k_1, k_2, k_3 \) and \( p_\psi \) be the momenta of incoming photon, incoming gluon, outgoing gluon and outgoing \( J/\psi \), respectively (see Fig. 1), \( \epsilon_1, \epsilon_2, \epsilon_3 \) and \( \epsilon_\psi \) be the polarization vectors, and \( a \) and \( b \) the colour indices of incoming and outgoing gluons. The photon gluon fusion matrix element reads:

\[
\mathcal{M}(\gamma g \rightarrow \psi g) = \text{tr} \left( \epsilon_1 (\not{p_\gamma} - \not{k}_1 + m_\gamma) \epsilon_2 \right) \times \epsilon_3 \left( -\not{k}_3 + m_\gamma \right) \epsilon_\psi J(\epsilon_\psi) \frac{1}{\sqrt{3}} \text{tr} (T^a T^b) \\
\times \left[ k_1^2 - 2 (p, k_1) \right]^{-1} \left[ k_3^2 + 2 (p, k_3) \right]^{-1} + 5 \text{ permutations of all gauge bosons.}
\] (5)

Here the projection operator [10] \( J(\epsilon_\psi) = \epsilon_\psi (\not{p_\psi} + m_\psi) / m_\psi \) guarantees the proper spin structure of \( c\bar{c} \) state, and the charmed quarks are assumed to each carry one half of the \( J/\psi \) momentum, \( p_\gamma = p_\psi = p_{\psi0} / 2, m_c = m_\psi / 2 \). The evaluation of trace in (5) is straightforward and is done using the algebraic manipulation system FORM [22].

The multiparticle phase space \( \Pi d^3 p_i / (2 E_i) \delta^4(\sum p_i - \sum p_{\text{out}}) \) is parametrized in terms of transverse momenta, rapidities and azimuthal angles:

\[
\frac{d^3 p_i}{2 E_i} = \frac{d^2 k_i}{2 \pi} d\phi_i d\eta_i. \\
\text{Let} \quad p_\gamma \quad \text{and} \quad p_{\psi} \quad \text{be the momenta of initial and scattered leptons,} \\
p_i \quad \text{the momentum of initial proton,} \\
s = (p_\gamma + p_{\psi})^2, \quad \eta_\gamma, \phi_\gamma, \eta_3 \quad \text{and} \quad \phi_3 \quad \text{the rapidity and the azimuthal angle of} \quad J/\psi \quad \text{and of accompanying gluon, respectively. Then, the fully differential cross section reads:}
\]

\[
d\sigma (e p \rightarrow e\psi X) = \frac{\alpha_s e_\gamma^2 |\Psi(0)|^2}{16 \pi s^2} \\
\times \frac{1}{4} \sum_{\text{spins}} \sum_{\text{colours}} \left| \mathcal{M}(\gamma g \rightarrow \psi g) \right|^2 \\
\times \frac{1}{x_2} \Phi \left( x_2, k_2^2, \mu^2 \right) dk_2^2 \frac{2}{2\pi} dy_3 d\phi_3 \\
\times \frac{1}{2\pi} d\phi_2 \frac{1}{2\pi} d\phi_2 \\
\] (6)

The phase space physical boundary is determined by the inequality

\[
G(\hat{s}, t, k_3^2, k_1^2, k_3^2, m_\psi^2) \leq 0.
\] (7)

where \( \hat{s} = (k_1 + k_2)^2 \), \( t = (k_1 - p_\psi)^2 \), and \( G \) is the standard kinematical function [23].

When calculating the spin average of the matrix element squared, we substitute the full lepton tensor for the photon polarization matrix:

\[
\overline{\epsilon_\gamma \epsilon_\gamma^*} = 4 \pi \alpha \left[ 8 p_\gamma^a p_e^a - 4 (p, k_1) g^{\mu\nu} \right] / (k_1^2) \\
(8)
\]

(including also the photon propagator factor and photon-lepton coupling). For the off-shell incoming gluon we take \( \epsilon_\gamma^\dagger \epsilon_\gamma^* = k_2^a k_2^a / k_2^2 \). The outgoing gluon is assumed on-shell, \( \sum \epsilon_\gamma^\dagger \epsilon_\gamma^* = -g^{\mu\nu} \). The \( J/\psi \) polarization vector \( \epsilon_\psi \) is defined as an explicit four-vector. In the frame with z-axis along the \( J/\psi \)
momentum, $p_\rho = (0, 0, |p_\rho|, E_\rho)$, it reads for different helicity states:

$$\epsilon_{\rho(h = \pm 1)} = (1, \pm i, 0, 0)/\sqrt{2},$$
$$\epsilon_{\rho(h = 0)} = (0, 0, E_\rho, |p_\rho|)/m_\rho. \quad (9)$$

The summation over colours yields a factor of 2/3.

For comparison, we also do calculations with on-shell initial gluons and/or photons. Then we replace the noncollinear gluon distribution $\Phi(x, k_T^2, \mu^2)/x_2$ with conventional gluon density $F_g(x_2, \mu^2)$ and drop the integration over the gluon transverse momentum and azimuthal angle. The on-shell photon is considered as a ‘parton’ in a lepton with its distribution function given by Weiszacker-Williams formula $F_\gamma(x_1) = \alpha/(2\pi) [1 + (1 - x_1)^2]/x_1 \ln(x/4m_e^2)$.

The photon and gluon momentum fractions $x_1$ and $x_2$ are calculated from the energy-momentum conservation laws in the light cone projections:

$$(k_1 + k_2)_{\perp} = 2E_x x_1$$
$$= m_\rho r \exp(y_\rho) + |k_{3T}| \exp(y_3),$$

$$(k_1 + k_2)_{\perp} = 2E_\rho x_2$$
$$= m_\rho r \exp(-y_\rho) + |k_{3T}| \exp(-y_3). \quad (10)$$

$m_{3T} = (m_\rho^2 + |p_{3T}|^2)^{1/2}$. The multidimensional integration in (6) has been performed by means of Monte Carlo technique, using the routine VEGAS [24].

4. Numerical results and discussion

The role of the photon and gluon virtualities may be seen in Fig. 2, where we show a comparison of three calculations made under the assumptions that two, or one, or none of the colliding gauge bosons are on mass shell. The results correspond to HERA conditions, i.e. electron proton collisions at the energy $\sqrt{s} = 300$ GeV, where we have applied the kinematical cuts $0.4 \leq z \leq 0.8$ in order to minimize the possible influence of other production mechanisms (e.g., fragmentation and/or colour octet contributions).

With respect to the inclusive $J/\psi$ leptonproduction cross section one observes only a small difference between the exact electromagnetic matrix element calculations and Weiszacker-Williams approxi-
As this intrinsic angle discussed in the context of di-jets photoproduction, the $J/\psi$ transverse momentum distribution gets much broader. However, the only kinematical effect yet cannot testify unambiguously for the validity of BFKL evolution. In fact, it can be rather easily imitated within some phenomenological models, such as PYTHIA [25], despite the latter is based upon collinear parton evolution and on-shell interaction matrix elements.

The situation becomes more conclusive as soon as the polarization variables are concerned. The off-shell photons and gluons do promptly manifest in the $J/\psi$ spin alignment. Only a very small fraction of $J/\psi$ particles can be produced in the helicity zero (i.e. 'longitudinal' polarization) state by massless gauge bosons (Fig. 2b). This property is totally determined by the photon gluon fusion matrix element structure.

The effects of initial gluon off-shellness may be, best of all, seen in the transverse momentum spectra, because the gluon virtuality is proportional to its transverse momentum: $m^2 = -k_T^2/(1-x)$. In contrast with the conventional (massless) parton model, semihard theory shows that the fraction of longitudinally polarised $J/\psi$ particles increases with $p_T$. A deviation from the parton model behaviour becomes well pronounced already from $p_T > 3$ GeV (Fig. 2c), and at $p_T > 6$ GeV the longitudinal polarization tends to be even dominant.

Qualitatively, the difference between the model predictions refers to different origins of the $J/\psi$ transverse momentum. In the case of conventional parton model, $J/\psi$ obtains its transverse momentum from the hard photon gluon interaction, while the contribution from primordial photon transverse motion is negligible (remind that the gluons have no transverse motion at all). On the contrary, within the semihard theory, the contribution from initial gluon transverse momentum is large. The distributions over other kinematical variables (such as the rapidity $y$ and the elasticity parameter $z$) do not exhibit strong dependence on the gluon virtuality and are not presented in this paper.

The distributions calculated in the absence of cuts on $z$ look very similar in their shape and only correspond to larger visible cross sections. This property can be understood from Fig. 3, which shows that the fraction of longitudinally polarized mesons remains almost constant over the whole $z$ range.

We point out that the degree of spin alignment may be measured experimentally since the different polarization states of $J/\psi$ result in significantly different angular distributions of the $J/\psi \rightarrow l^+l^-$ decay leptons:

$$d\Gamma_{h=\pm 1}/d\cos \Theta \sim 1 + \cos^2 \Theta,$$

$$d\Gamma_{h=0}/d\cos \Theta \sim 1 - \cos^2 \Theta.$$  \hspace{1cm} (11)

(Here $\Theta$ is the angle between the lepton and $J/\psi$ directions, measured in the $J/\psi$ rest frame, and the subscript $h$ stands for the $J/\psi$ helicity). In general, a mixed polarization sample would realize in the angular distribution:

$$d\Gamma/d\cos \Theta \sim 1 + a \cos^2 \Theta,$$

$$a = (1 - 3h_0)/(1 + h_0),$$  \hspace{1cm} (12)

with $h_0$ being the fraction of longitudinally polarized $J/\psi$ mesons.

\[ \sigma_{h=0}/\sigma \]

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fraction.png}
\caption{The fraction of longitudinally polarized mesons versus the inelasticity parameter $z = E_\phi/E_T$.}
\end{figure}

\footnote{Actually, PYTHIA gives some intrinsic $k_T$ to the incoming partons. This can account for the $p_T$ broadening effect and the deviations from ideal back-to-back configurations in the azimuthal angle discussed in [8] in the context of di-jets photoproduction. As this intrinsic $k_T$ develops from partonic branching cascades, it may be said that part of the relevant physics is included in PYTHIA. However, the way it is organized is phenomenological rather than theoretical, because the $k_T$ evolution algorithm does not rely upon BFKL or any other consistent theory. Of course, PYTHIA does not predict an enhanced rate for longitudinally polarized mesons (see below) because it uses the on-shell matrix elements only.}
Finally, the most important theoretical uncertainties, such as the ones related to gluon densities and the strong coupling constant (see discussion in [26] on the proper choice of $\mu^2$ in $\alpha_s(\mu^2/\Lambda^2)$), may only affect the production rate $\sigma(ep \rightarrow J/\psi X)$ but not the shape of the lepton angular distribution (12). Consequently, the latter may serve as a perfect indicator for the initial gluon off-shellness.

5. Conclusions

Gluon virtuality connected with its transverse momentum is one of the inherent properties of non-collinear BFKL parton evolution theory. Compared to traditional (collinear) parton model, gluons are characterized by a different spin density matrix. The latter is found to affect the polarization of $J/\psi$ particles produced in $ep$ collisions via photon gluon fusion subprocess. The effect is best pronounced at large $J/\psi$ transverse momenta, and may be detected experimentally by measuring the $J/\psi \rightarrow l^- l^+$ decay lepton angular distributions. We recommend the above process as a direct probe of the gluon virtuality, which can eventually testify for the validity of noncollinear parton evolution.

Acknowledgements

The author conveys his thanks to Hannes Jung and Nikolai Zotov for their encouraging interest and useful discussions.

References

Leading proton spectrum from DIS at HERA

A. Szczurek a, N.N. Nikolaev b, J. Speth b

a Institute of Nuclear Physics, PL-31-342 Cracow, Poland
b Institut für Kernphysik, Forschungszentrum Jülich, D-52425 Jülich, Germany

Received 28 January 1998
Editor: P.V. Landshoff

Abstract

The QCD hardness scale for secondary particles (h) production in semi-inclusive deep inelastic scattering (DIS), $ep \rightarrow e'Xh$, gradually decreases from $Q^2$, the photon virtuality which determines the hard scale in the virtual photon (current) fragmentation region to a soft, hadronic, scale in the proton fragmentation region. This suggests similarity of the inclusive spectra of leading protons and neutrons, $h=p,n$, in high energy hadron-proton and virtual photon-proton collisions. We explore this similarity extending to the DIS regime the nonperturbative peripheral mechanisms of inelastic scattering traditionally used in hadronic interactions to explain fast nucleons production. While the production of leading neutrons is known to be exhausted by DIS off charged pions, the production of leading protons by DIS off neutral pions must be supplemented by a substantial contribution from isoscalar reggeon $f_0$ exchange extrapolated down to moderate values of $x$. We comment on the $x$ and $Q^2$ dependence of leading proton production as a probe of a universal pattern of the $x,Q^2$ evolution of the nucleon and meson (reggeon) structure functions at small $x$. © 1998 Elsevier Science B.V. All rights reserved.

PACS: 12.40.Nn; 13.60.Hb; 13.85.Ni

In deep inelastic $ep$ scattering, according to the standard QCD description of hadronization, the proton fragmentation region is very different from the current and/or photon fragmentation region. Namely, the virtuality of partons in generalized ladder diagrams gradually decreases from the hard scale $Q^2$ of the struck parton in the current fragmentation region to the soft, hadronic, scale of the parent parton in the proton fragmentation region.

Although until quite recently experimental data on proton fragmentation were scarce [1], presently the ZEUS and H1 leading proton spectrometers (LPS) and forward neutron calorimeters (FNC) are operational and are amassing data on leading proton and neutron production [3,4].

Whereas popular Monte Carlo implementations of perturbative QCD (Ariadne [5], Herwig [6] and others) are very successful in the photon fragmentation region (for a recent review see [7]), a purely perturbative description of the proton fragmentation region is not yet possible and the current versions of Monte Carlo hadronization models underestimate the yield

Before the ZEUS/H1 LPS era, leading protons in DIS have been studied only in the fixed-target bubble-chamber neutrino experiments [1]. Only large values of the Bjorken $x$ were accessible, leading to a strong kinematical bias in the leading proton spectra [2]. The small-$x$ data from HERA are free of this kinematical bias.
of fast secondary nucleons [8]. Traditionally, leading proton production in inelastic collisions is modeled via nonperturbative peripheral interactions. Such peripheral models were quite successful when applied to hadronic collisions (for a review see [9]). A well known example of this type of processes is the interaction of projectiles with pions from the chiral mesonic substructure of the proton. We recall that in hadronic reactions, the pion exchange mechanism with absorption exhausts the cross section for the hadronic reactions, the pion exchange mechanism ; c protons produced as decay products of fast baryon resonances of which the Δ production via pion-exchange is a typical, and predominant, source; d reggeon heavy meson (reggeon R) exchange contribution (predominantly the isoscalar reggeon, \( R = f_{0} \), exchange). A preliminary evaluation of the first two mechanisms has been done in [16]. The background from pion and reggeon exchange to the dominant pomeron exchange at \( z \geq 0.95 \) has been discussed recently in [20–23], and we partially use the results of these works. Diffraction excitation of the proton into high-mass states also contributes to leading proton production, and we shall comment on this small contribution following the considerations in [24].

Under approximations to be specified below, the contributions of all four mechanisms to the semi-inclusive structure function can be written in the factorized form (\( i = P, \pi^{0}, p, \pi \Delta, f_{0} \)):

\[
F_{2}(z, t, Q^{2}) = \sum_{i} f_{i}(z, t) \cdot F_{i}(\beta, Q^{2}) ,
\]

(2)

where \( F_{i}(\beta, Q^{2}) \) is the structure function of the exchanged object (pion, pomeron, reggeon), \( f_{i}(z, t) \)

![Fig. 1. Peripheral mechanisms of leading proton production.](image-url)

For related work on electro-magnetic properties of nucleons and W-boson and jet production in nucleon-nucleon collisions see [13,14].

used in hadronic reactions to lepton DIS at HERA, we consider four mechanisms of leading proton production (Fig. 1): a) diffractive production of protons (pomeron \( P \) exchange) which dominates at \( z \to 1 \), and constitutes a background to fragmentation in non-diffractive DIS at \( z \leq 0.9-0.95 \); b) spectator protons from the fragmentation of the \( \pi N \) Fock state of the physical proton produced by DIS off virtual \( \pi^{0} \) (pion-exchange mechanism); c) protons produced as decay products of fast baryon resonances of which the Δ production via pion-exchange is a typical, and predominant, source; d) reggeon heavy meson (reggeon \( R \) ) exchange contribution (predominantly the isoscalar reggeon, \( R = f_{0} \), exchange). A preliminary evaluation of the first two mechanisms has been done in [16]. The background from pion and reggeon exchange to the dominant pomeron exchange at \( z \geq 0.95 \) has been discussed recently in [20–23], and we partially use the results of these works. Diffraction excitation of the proton into high-mass states also contributes to leading proton production, and we shall comment on this small contribution following the considerations in [24].
is its flux factor and $\beta$ is the Bjorken variable for DIS off the exchanged object.

We start our discussion with the pion exchange mechanism. In this case $F^{\prime}_{2}(\beta, Q^{2})$ is the structure function of the physical pion and the flux factor is given by

$$f_{\pi^{+}}(x_{\pi}, t) = \frac{8s_{\pi^{+}}^{2}}{16\pi^{2}}(1-z)\frac{(-t)|F_{\pi^{+}}(z,t)|^{2}}{(t-m_{\pi}^{2})^{2}}.$$  \hspace{1cm} (3)

Strictly speaking, Eq. (3) holds in the plane wave impulse approximation. A recent analysis [10] has shown that absorption corrections to pion exchange in DIS are small and can be neglected for the purposes of the present analysis. Also, the off-mass shell extrapolations are marginal and the on-mass shell pion structure function can be used. Important consistency check is provided by the simultaneous description of the hadronic leading nucleon data. The results for DIS in the interesting region of $0.6 \leq z \leq 0.9$ only marginally depend on whether the light-cone or Regge parameterization of $|F_{\pi N}(z,t)|^{2}$ are used (for a detailed discussion concerning the choice of the form factor see Refs. [10,12]).

Production of fast $\Delta^{+}$'s is also known to be dominated by pion exchange. For $\Delta^{++}$ production the flux factor is given by

$$f_{\Delta^{++}}(z,t) = \frac{2s_{\pi^{+}}^{2}4m_{\Delta}^{2}}{16\pi^{2}}(1-z)\times \frac{((m_{\Delta}+m_{N})^{2}-t)((m_{\Delta}-m_{N})^{2}-t)|F_{\Delta^{++}}(z,t)|^{2}}{6m_{\Delta}^{3}(t-m_{\pi}^{2})}.$$  \hspace{1cm} (4)

Contributions from $\Delta^{+}$ and $\Delta^{0}$ production can be included using the familiar isospin relations (see for instance [12,24]). In the simplest one-pion exchange approximation, the polarization state of the produced $\Delta$'s is such that the $\Delta \to \pi N$ the decay angular distribution in the Gottfried-Jackson (t-channel) frame equals

$$w(\theta_{f},\phi_{FY}) = 1/4 \cdot (1 + 3\cos^{2}\theta_{f}) \cdot Y_{00}(\phi_{FY}),$$

where $\theta_{f}$ and $\phi_{FY}$ are the so-called Jackson and Treiman-Yang angles, respectively. Absorptive correction modify slightly this simple form [25], but these corrections can be neglected for the purposes of the present analysis, since both the $z-$ and $t-$ spectra of decay protons only weakly depend on the decay angular distributions.

For the diffractive $e+p \to e^{'}+p^{'}+X$ reaction, our semi-inclusive structure function coincides with the pomeron component of the diffractive structure function. $F^{D}_{2}(P_{N};z,t,\beta, Q^{2}) = F^{D}_{2}(\beta, Q^{2})$. At $z \leq 0.9$, diffractive DIS is a small background to non-diffractive DIS and a somewhat simplified description is justified. Since the ZEUS data have been taken at $|\slant t| \leq 10^{-3}$ and $z \leq 0.9$ then $\beta$ is quite small, $\beta \leq 2 \cdot 10^{-3} - 10^{-2}$ and it has been argued [26,27] that at such a small $\beta$ one expects the factorization

$$F^{D}_{2}(\beta, Q^{2}) = f_{P}(z,t) \cdot F_{2}^{P}(\beta, Q^{2}).$$  \hspace{1cm} (5)

The normalization of the pomeron flux factor $f_{P}(z,t)$ and the pomeron structure function $F_{2}^{P}(\beta, Q^{2})$ is a matter of convention, and only the product of the two is well defined. To be specific, we use the triple-Regge parameterization for the flux factor

$$f_{P}(z,t) = \frac{1}{8\pi^{2}(1-z)}(1-z)^{2(1-\alpha_{P}(t))}G_{P}(t),$$  \hspace{1cm} (6)

where $G_{P}(t) = G_{0}\exp(B_{p}t)$ with $G_{0} = 21.2$ mb [21–23] from the Regge decomposition of the NN total cross sections [28] and $B_{P} = 3.8$ GeV$^{-2}$ according to the triple-Regge analysis of hadronic diffraction scattering [9,29–31]. For $z \leq 0.9$, the specific Regge effects comes from $(1-z)^{2(1-\alpha_{P}(t))}$ and from the $t$-dependence of the pomeron trajectory $\alpha_{P}(t)$ are marginal and, besides the standard factor $\frac{1}{t}$, the main $z$ dependence of the flux comes from the kinematical boundary $|t| \geq 1|t|_{max} = \frac{\sqrt{Q^{2} - 4m_{\pi}^{2}}}{2}$ in the form factor $G(t)$. In principle $F^{D}_{2}(\beta, Q^{2})$ can be derived from experimental data on diffractive DIS, but currently for $\beta \leq 2 \cdot 10^{-3} - 10^{-2}$ the pomeron structure function stays basically unknown. It has been argued [26], that at small $\beta$ the conventional DGLAP evolution holds for the pomeron structure function giving a $\beta$ and $Q^{2}$ dependence of $F^{P}_{2}(\beta, Q^{2})$ similar to that of $F^{P}_{2}(\beta, Q^{2})$ (see for instance [27]). On the other hand, the triple-pomeron formula with soft pomerons gives the scaling prediction $F^{P}_{2}(\beta, Q^{2}) = C_{P} \beta^{-0.08}$ (the normalization $C_{P}$...
= 0.026 has been fitted [23] to the H1 experimental data [20]. We use these two models to check the sensitivity of the leading proton spectra to the evolution in $\beta$ and $Q^2$.

The reggeon exchange is an important ingredient of the triple-Regge phenomenology of hadronic diffraction, although its strength is not very well known [9,29,30]. The triple-Regge parameterization for the reggeon flux is

$$f_{R}(z,t) = \frac{1}{8\pi} (1-z)^{1-2a_{R}(t)} G_{R}(t),$$

(7)

where $a_{R}(t)$ is the reggeon trajectory. We take $G_{R}(0) = G_{R}(0) \cdot \exp(B_{R} t)$, where for the dominant $f_{0}$-exchange $G_{R}(0) = 76$ mb [21,23] and $B_{R} = 4$ GeV$^{-2}$, which is consistent with the data on leading proton production in $pp$ collisions [32]. Triple-Regge considerations in conjunction with fits to the NN final state have suggested the suppression factor

where $G_{R}$ is the reggeon trajectory. We take $G_{R}(0) = G_{R}(0) \cdot \exp(B_{R} t)$, where for the dominant $f_{0}$-exchange $G_{R}(0) = 76$ mb [21,23] and $B_{R} = 4$ GeV$^{-2}$, which is consistent with the data on leading proton production in $pp$ collisions [32].

The single particle inclusive $(z,t)$-spectrum of protons is defined as

$$R(z,t,x,Q^2) = F^{pp}(z,t,x,Q^2)/F_{2p}(x,Q^2).$$

A fully differential study of $R(z,t,x,Q^2)$ is not yet possible with the limited statistics of the preliminary ZEUS data [3]. The data were collected within the following experimental cuts $\Omega_{\exp}: 0.6 < z < 0.9, |t|_{\exp} < 0.5$ GeV$^{-2}, 10^{-4} < x < 10^{-3}$ and $4 < Q^2 < Q_{\max}^2$, where $Q_{\max}^2$ is the maximal kinematically attainable $Q^2$. Within these cuts the fraction of events with leading proton is given by

$$R_{\exp} = \sum_{i} R_{\exp}^{i} = \sum_{i} \frac{\Delta \sigma^{i}(\Omega_{\exp})}{\Delta \sigma^{tot}(\Omega_{\exp})},$$

where

$$\Delta \sigma^{i}(\Omega_{\exp}) = \int_{t_{\min}}^{t_{\max}} dx \int_{x_{\min}}^{x_{\max}} dx \int_{Q_{\min}^2}^{Q_{\max}^2} dQ^2 \frac{d\sigma^{i}}{dx dQ^2 dt},$$

$$\Delta \sigma^{tot}(\Omega_{\exp}) = \int_{x_{\min}}^{x_{\max}} dx \int_{Q_{\min}^2}^{Q_{\max}^2} dQ^2 \frac{d\sigma^{tot}}{dx dQ^2},$$

and the subscript $i$ stands for one of the mechanisms shown in Fig. 1.

As emphasized above, the pion, pomeron and reggeon structure functions are unknown in the $\beta$ region considered in our present analysis. For a reference evaluation of $R_{\exp}$, we take the GRV parameterization for the $(\beta,Q^2)$ evolution of the pion structure function [33], the flux of pions evaluated in the light-cone model for the chiral structure of the nucleon [12] (the Regge parameterization leads to very similar result [10]) and the triple-Regge model parameterization $F_{2}^{p}(\beta,Q^2) = 0.026 \beta^{-0.08}$ described above. For the proton structure function, which enters the evaluation of the denominator (10) of the ratio (8), one can use any convenient fit to the HERA data. In the present analysis we take the GRV parameterization [34]. The practical calculations have been performed with a Monte Carlo implementation of the above formalism and for the ZEUS kinematical cuts. As a result of our analysis, we find $R_{\exp}^{\beta}(\Omega_{\exp}) = 2\%$, $R_{\exp}^{p}(\Omega_{\exp}) = 0.9\%$ and the tail of the pomeron exchange contribution [22] gives $R_{\exp}^{p}(\Omega_{\exp}) = 1.2\%$, so that $R_{1+2+3}^{p}(\Omega_{\exp}) = \frac{1}{1+2+3} R_{\exp}^{p}(\Omega_{\exp}) = \frac{1}{1+2+3} \times 1.2\%.$
4.05%. Note that QCD hadronization models (Ariadne, Herwig) were never meant to describe the nonperturbative proton fragmentation region; for example Ariadne [5] gives \( R(\Omega_{\text{exp}}) \) in the per mill range and a similar under-prediction for the production of leading neutrons [8]. From the comparison with the preliminary ZEUS experimental result, \( R_{\text{ZEUS}}^{\text{exp}} = 9.2 \pm 1.7 \% \) (stat. only) [3], we conclude that about 5% of the missing strength must be attributed to the reggeon exchange. In the triple-Regge model, \( F_2^R(\beta,Q^2) = C_R \beta^{-0.08} \), this requires \( C_R = 0.12 \) within a factor of 1.5 uncertainty.

Having fixed the strength of the reggeon exchange, we can see better the importance of different \( z \)-dependence of the ratio \( R_{\text{exp}}(z) \) defined for the experimental \((t,x,Q^2)\) range as shown in Fig. 2. The importance of the reggeon exchange is obvious. With the set of parameters specified above, the reggeon contribution makes \( R_{\text{reg}}(z) \) approximately flat at \( z \leq 0.9 \), in close similarity to a flat \( z \)-spectrum of leading protons in hadronic interactions [32]. The preliminary H1 results are also consistent with the flat \( z \)-spectrum [4].

The \( z \)-spectrum of leading nucleons from diffraction double dissociation (DD) has been studied in [24]. It can be isolated experimentally by the rapidity gap (GAPCUT) selection method [3]. An extension of the analysis [24] to leading protons shows that \( \sim 70\% \) of DD events have leading protons, mainly produced by excitations of the \( N\pi\pi \) and high mass continuum states. Roughly \( \sim 50\% \) of final state protons have \( z > 0.6 \). Since DD constitutes \( \sim 2\% \) of the DIS events, only \( \sim 0.7\% \) have a leading proton with \( z > 0.6 \) generated by this mechanism. DD is therefore a small, \( f(\text{GAPCUT}) = R_{\text{DD}}(\Omega_{\text{exp}})/R_{\exp} \sim 7\% \), background to the dominant non-diffractive production mechanism, in good agreement with the ZEUS findings. The LEPTO6.5 ‘soft color interaction’ model [35] which, unlike Ariadne and/or Herwig, is supposed to describe all aspects of DIS including leading proton production, over-predicts the fraction of GAPCUT events: \( f(\text{GAPCUT}) = 20-30\% \). The observed leading proton \( z \)-distribution of the GAPCUT sample [3] is also consistent with the \( z \) spectrum of protons generated in proton dissociation into \( N\pi\pi \) and continuum states as shown in Fig. 2 of [24]. It can readily be included in the analysis of higher precision data.

This evaluation of the reggeon exchange from fragmentation into protons at \( z \leq 0.9 \) is consistent within a factor of 2 with estimates of the reggeon background to pomeron exchange in the diffractive region of \( z \geq 0.95 \) [23]. A caveat in comparing these two extreme regions is the possible reggeon-pomeron interference contribution \( \sim \frac{1}{\log x} \), which can be substantial in the diffractive domain and small in the fragmentation region \( z \leq 0.9 \). A combined analysis of high precision fragmentation and diffractive data would be the best way to fix the reggeon-pomeron interference contribution, but such an involved phenomenology is not warranted with the presently available data. Note also that the ZEUS data are preliminary and lacking the evaluation of systematic errors.

In Fig. 3 we show the slope \( b(z) \) of the \( t \)-distributions defined in the experimental range of \((x,Q^2)\) \((R(z,t) \propto \exp(b(z)t))\). The slope of the reggeon trajectory is large, \( \alpha'_R = 0.9 \) GeV\(^{-2} \), and for pure reggeon exchange contribution quite a substantial rise of the slope is expected as \( z \) increases: \( b_R(z) = B_R + 2 \alpha'_R \log \frac{1}{1 - z} \). Similar growth of the slope is expected also for the pion exchange contribution. The increase of the slope at large \( z \) is tamed by the
small diffraction slope of the pomeron contribution. The parameter $B_R$ is poorly known and the leading proton spectrum offers the best possibility for its determination. In Fig. 3 we show the results for $b(z)$ obtained with $B_R = 4 \text{ GeV}^{-2}$. They are very close to the slope of the $t$-dependence for leading protons observed in $pp$ collisions [32]. The results shown in Fig. 3 are consistent with the preliminary ZEUS data [3] which give the slope $b(z)$ in the range $4-8 \text{ GeV}^{-2}$.

The $(x,Q^2)$-dependence of different mechanisms is controlled by the ratios $r(x,Q^2) = F_{2p}(x,Q^2)/F_{2b}(x,Q^2)$. In the scenario with the scaling soft pomeron/reggeon structure functions, $r(x,Q^2)$ decreases with rising $Q^2$ and/or decreasing $x$, because of the scaling violations and steep $x$-dependence of the proton structure function $F_{2p}(x,Q^2)$. Fig. 4 shows that in such a scaling scenario (no QCD evolution for $F_{2p,R}(\beta,Q^2)$), one would expect significant dependence of the leading proton production on both $x$ and $Q^2$. On the other hand, if $F_{2p,R}(\beta,Q^2)$ satisfies the conventional $(\beta,Q^2)$ evolution at small $\beta$ [26], one would expect very weak ($x,Q^2$) dependence of the leading proton spectrum. This stays true also in the real photoproduction limit. In the present analysis we model the evolution effects by taking $F_{2p,R}^z(\beta,Q^2) = \lambda_{p,R} F_{2p}^z(\beta,Q^2)$ with the GRV pion structure function. We adjust $\lambda_p = 0.2$ and $\lambda_r = 0.5$ as it was evaluated in [21] so that we reproduce the same $R_{exp}$ as with the scaling (no evolution) scenario within the ZEUS kinematical cuts.

In the conventional evolution scenario we indeed find a very weak $x$ and $Q^2$ dependence of the leading proton spectra. The preliminary ZEUS data [3] better agree with this scenario, although the error bars are still rather large. The preliminary H1 data [4] on the $t$-integrated $F_{2p}^z(z,t,x,Q^2)$ also support the conventional evolution scenario.

We conclude that the salient features of fragmentation into leading protons can be understood quantitatively.

![Fig. 3](image1.png)

Fig. 3. The slope of the $t$-distributions predicted by the model for $B_R = 4 \text{ GeV}^{-2}$.

![Fig. 4](image2.png)

Fig. 4. (a) The sensitivity of the fraction of DIS events containing leading protons within ZEUS cuts to the $Q^2$ evolution effects for the two different scenarios for the pomeron and reggeon structure function: the curves marked by filled circles are for the no-evolution soft pomeron model, the unmarked curves show the results for the conventional QCD evolution scenario modeled by the GRV pion structure function. The legend of curves is the same as in Fig. 2. (b) The same as Fig. 4a, but for the $x$-dependence of the fraction of DIS events containing leading protons within ZEUS cuts. The legend of curves is the same as in Fig. 2.
tatively in terms of peripheral mechanisms extended to the DIS regime. The experimentally observed similarity of the leading proton spectra in $pp$ collisions and $ep$ DIS is a natural consequence of these mechanisms. We emphasize that our approach has the capability of a unified description of diffractive DIS at $z \gtrsim 0.9$ and of fragmentation into protons in non-diffractive DIS. Of the four sources of leading protons pion exchange can be experimentally determined using neutron tagged DIS. Experimental confirmation of our estimate for this process will lend strong support also for our evaluation of the $\Delta$ contribution. The pomeron exchange background can be inferred from diffractive DIS. Finally, the reggeon exchange mechanism of fragmentation can also be tested in diffractive DIS. The combined analysis of the high precision leading proton data and diffractive DIS data makes possible a determination of the reggeon-pomeron interference effects, which has not been accomplished with the hadronic diffraction data [9, 29, 30].

A comparison of the soft pomeron no-evolution (unrealistic though it is) and conventional evolution scenarios for the reggeon structure function shows that the high precision leading proton spectrum offers an interesting test of the universality of the QCD evolution properties of structure functions at small $x$. (For a related discussion of the pion exchange mechanism within Veneziano’s fracture function [36] context see [37].)

We conclude with the comment that similar fragmentation mechanisms may be at work also at smaller $z$, where DIS on the multi-pion Fock states of the physical nucleon, $(n\pi)^N, (n\pi)^-\Delta, (n\pi)^N^+, (n\pi)^\Delta^+$, provides a natural mechanism for slowing down secondary protons. In the spirit of the above discussion, the weak $xQ^2$ dependence must hold also for slower protons. Similar arguments hold for the fragmentation of protons into hyperons $(A, \Sigma, \ldots)$.

Acknowledgements

At all stages of this work we benefitted greatly from the continuous advice of, and enlightening discussions with, N. Cartiglia of the ZEUS Collaboration. One of us (A.S.) is indebted to K. Golec-Biernat for interesting discussion. This work was supported partly by the German-Polish exchange program, grant No. POL-81-97 and by the INTAS Grant ’97-0597.

References