Baryon Number Violation at High Energy in the Standard Model: Fact or Fiction?*

Michael Dine†
Santa Cruz Institute for Particle Physics
University of California, Santa Cruz, CA 95064

Abstract

In the standard model, baryon and lepton number are not strictly conserved, due to anomalies. It has long been known that at low energies, the resulting baryon number violating amplitudes are extraordinarily small. A number of authors have suggested that at high temperatures or energies, baryon number violating effects should be enhanced. We give simple arguments that while baryon number violation is indeed large at high temperatures, there is no such enhancement at high energies.

* Work supported in part by the U.S. Department of Energy.
† on leave of absence from the City College of CUNY

We are all well aware that baryon number is conserved in nature to an extraordinary degree of accuracy. On the other hand, grand unified theories and string theory predict that there should be some small violation of baryon number in the microscopic equations of motion, and the existence of an asymmetry between antimatter and matter strongly suggests that such a violation exists.[1] Even in the standard model, baryon number is not strictly conserved. At very low energies, the violations of baryon (and lepton) number predicted by the model were computed by 't Hooft.[2] Fortunately, they are unobservably small. This is because baryon number violation in the standard model is associated with tunneling through a large barrier. This tunneling can be described by familiar semiclassical methods and leads to exponentially small amplitudes. At sufficiently high energies and temperatures, however, one might wonder whether these effects could be enhanced, since enough energy would be available to pass over the barrier without tunneling. Perhaps the first concrete suggestion along these lines was due to Klinkhammer and Manton.[3] These authors found the configuration corresponding to the top of the potential barrier (a static field configuration know as the "sphaleron") and computed the barrier height. They speculated (as we will see correctly) that at high temperature, it would be relatively easy to pass over the barrier, but that at high energies it would be extremely difficult.

Subsequently, Kuzmin, Rubakov and Shaposnikov (KRS) elevated the speculations concerning high temperature baryon number violation to a set of serious calculations.[4] Further works verified and fleshed out this picture.[5,6] KRS also speculated that there might be an enhancement of the cross section for high energy scattering, and/or for decay of heavy particles. The idea of high energy enhancement has received support recently from a set of instanton calculations by Ringwald[7] and Espinosa.[8] These authors discovered that, at least for some range of energy, the total cross section grows exponentially with energy from its infinitesmal low energy value. McLerran, Vainshtein, and Voloshin[9] have argued that this exponential growth persists until the cross section saturates the unitarity bound, at energies possibly as small as 10 TeV! In these high energy activated tunneling processes, a pair of quarks, for example, would scatter producing three leptons and another 7 quarks, with a net violation of lepton number by -3 units and baryon number by 3 units. Perhaps even more striking, however, would be the production of a huge number [of order $\frac{1}{\alpha_2}$, where α_2 is the SU(2) gauge coupling] of W's, Z's and Higgs bosons.

In this lecture, we will perform a very simple calculation, directly in Minkowski space, which reproduces the results of Ringwald and Espinosa.[10] We will see that their answer can indeed be understood as resulting, in part, from a reduction with energy in the penalty one pays for tunneling over the barrier. However, we will also see that there is a suppression, coming from the difficulty in coupling to the mode which describes motion over the barrier. At energies low compared to the barrier height, the enhancement of the tunneling factor "wins," and the cross section grows exponentially rapidly, in precisely the fashion given by the instanton calculations. However, this cannot persist indefinitely; the cross section remains exponentially small at all energies.

In order to understand the issues involved, it will be helpful to review the question of baryon number violation at low energies in the standard model. At the classical level, the lagrangian of the theory preseves baryon number and the separate lepton numbers. Indeed, one of the elegant features of the standard model is that there is no dimension four (renormalizable) operator one may add to it which violates these quantum numbers. At the quantum level, however, these symmetries are violated. This breakdown of symmetry is associated with the phenomenon of "anomalies," familiar from $\pi^0 \to 2\gamma$ decay. Before considering the full standard model, consider first a simpler theory: SU(2) gauge theory with a single *massless* Dirac doublet. The Lagrangian density is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + i\bar{\psi}\gamma^\mu\partial_\mu\psi \quad . \tag{1}$$

This theory possesses, at the classical level, two global $U(1)$ symmetries,

$$\psi \to e^{i\alpha}\psi \qquad\qquad \psi \to e^{i\alpha\gamma_5}\psi \tag{2}$$

with corresponding currents

$$j_V^\mu = \bar{\psi}\gamma^\mu\psi \qquad\qquad j_5^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi \quad . \tag{3}$$

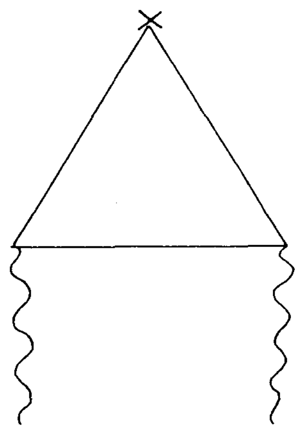


Figure 1. Triangle diagram gives rise to anomaly in axial current. Solid lines denote fermions, wavy lines are gauge bosons; X denotes insertion of current.

Quantum field theories are rather singular objects, and at the quantum level, it is not possible to enforce conservation of both the vector and axial currents. The triangle diagram of fig. 1, with an insertion of the axial current, j_5^μ, at one vertex, and gauge bosons at the other two vertices, is badly behaved in the ultraviolet, and care is required in its definition. One usually defines the theory so that the vector current is strictly conserved (in general, for consistency, one must define the theory so that any gauged currents are conserved), but the axial current is anomalous. The divergence of the axial current is given by

$$\partial_\mu j^{\mu 5} = \frac{g^2}{32\pi^2}\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} \tag{4}$$

$$= \frac{g^2}{16\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu} \quad .$$

One might think that having uncovered this anomaly, the axial charge,

$$Q_A = \int d^3x j^{05}$$

would simply not be conserved. However, the situation is more subtle, for another conserved current exists. By straightforward algebra, one can show that $F\tilde{F}$ is a total divergence,

$$F\tilde{F} = \partial_\mu K^\mu \tag{5}$$

where

$$K^\mu = \epsilon_{\mu\nu\rho\sigma}Tr[A_\nu\partial_\rho A_\sigma + \frac{2}{3}A_\nu A_\rho A_\sigma] \, .$$

Thus one can define a conserved current, $\tilde{j}^\mu = j^{\mu\,5} - \frac{g^2}{16\pi^2}K^\mu$, which obeys $\partial_\mu \tilde{j}^\mu = 0$. On the other hand, $\tilde{j}^\mu$ is not gauge invariant. In particular, it is not always true that the potentials fall fast enough at infinity to allow one to neglect surface terms (field strengths and similar gauge-invariant objects do always fall rapidly to zero at infinity). The problem, which will be discussed in more detail below, is most easily phrased in the language of the Feynman path integral. There one sums over all possible histories of the gauge (and other) fields, weighted by e^{iS}. In other words, one sums over all possible classical field configurations (not necessarily solutions to the classical equations of motion). It turns out that there are important classical configurations, known as instantons, for which

$$\frac{g^2}{16\pi^2}\int d^4x F\tilde{F} = n$$

so that $\Delta Q_5 = n$. 't Hooft showed that in this theory, these instanton configurations lead to an effective interaction for the massless fermions,

$$\mathcal{L}_{inst} = ce^{-\frac{8\pi}{g^2}}\bar{\psi}\psi \ . \tag{6}$$

Here c is a numerical constant. This interaction explicitly violates the axial symmetry. However, the coefficient is exponentially small for small coupling.

The generalization of all of this to the standard model is straightforward. Denote the left-handed quark and lepton doublets by Q_i and L_i, respectively, where the subscript is a generation index. Written as four component spinors, these particles couple to W bosons with the usual $\frac{1}{2}(1-\gamma_5)$ coupling. The right handed singlets can be described in terms of their left handed antiparticle counterparts, u_i^c, d_i^c and e_i^c. These fields do not couple to the $SU(2)$ gauge fields. Classically, baryon number conservation arises due to the symmetry of the lagrangian under the phase rotations

$$Q_i \rightarrow e^{i\frac{\alpha}{3}} Q_i \qquad (u_i^c, d_i^c) \rightarrow e^{-i\frac{\alpha}{3}} (u_i^c, d_i^c) \ . \tag{7}$$

Correspondingly, the baryon number current is

$$j_B^\mu = \frac{1}{3} \left[\sum_i \bar{Q} \frac{1}{2} (1 - \gamma_5) \gamma^\mu Q - \bar{u}_i^c \frac{1}{2} (1 - \gamma_5) u_i^c \right] \ . \tag{8}$$

Now, consider the triangle diagram involving the current, j_B^μ, and two $SU(2)$ gauge bosons. Only the Q fields appear in this diagram, since only they couple to the gauge bosons. As a result of the γ_5's appearing here, the diagram is anomalous. Similar remarks apply to lepton number. However, it is easy to see that $B - L$ *is* conserved. In the case of the standard model, 't Hooft calculated the leading baryon number violating term in the low energy effective lagrangian for the fermions due to instantons. For three generations (making, for simplicity the drastic approximation of neglecting the top quark mass), this term takes the form:

$$\mathcal{L}_B = c \left(\frac{1}{M_W} \right)^{14} e^{-\frac{2\pi}{\alpha_2}} Q_1^a Q_1^b Q_1^c \ldots Q_3^a Q_3^b Q_3^c L_1 L_2 L_3 \ . \tag{9}$$

Here the subscripts on the quark and lepton fields represent generation indices; the superscripts represent color indices, and we have suppressed $SU(2)$ indices.All of these indices must be contracted in a suitable way, but the details will not be important to us.

While interesting, this result, as it stands, is only a theoretical curiosity. The exponential factor is a number of order 10^{-70}, so the corresponding cross section for such processes is extraordinarily small. On the other hand, this effective action is relevant only at energies small compared to the W, Z, and Higgs boson masses. At these low energies, the effective action is obtained by first solving for the dynamics of the heavy fields in terms of the light fields; this

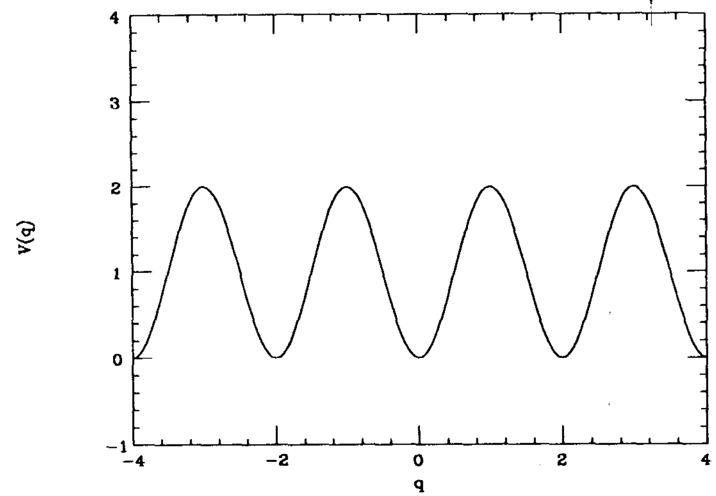


Figure 2. Schematic drawing of vacuum structure of a non-abelian gauge theory.

is known as "integrating out" the heavy fields. At higher energy scales or at very high temperatures, we must study the dynamics of the complete system. In particular, it is not a priori obvious that at extremely high energies, baryon violating processes will be so drastically suppressed.

In order to address this question, it is necessary to understand why the amplitudes at low energies are so small. Consider first a pure gauge theory (i.e., a theory with no fermions, or with only massive fermions with vector-like couplings). An Abelian gauge theory, like QED, has a very simple structure. Once one has made a suitable gauge choice, the ground state wave function is simply a Gaussian centered about the classical zero energy configuration, $A_\mu = 0$. This is not the case in a non-Abelian theory. Here, the classical condition for vanishing energy, $F_{\mu\nu}^a = 0$, has, loosely speaking, an infinite set of solutions even in a fixed gauge. Correspondingly, there is a large set of degenerate vacua, indicated schematically in fig.2. These vacua are labeled by an integer, N_{CS}, referred to as the Chern-Simons number, and are separated from one another by a barrier.* In a suitable gauge, the Chern-Simons number is related to the current, K^μ we encountered above through

$$N_{CS} = \int d^3x\; K^0 \;. \tag{10}$$

The figure is schematic, since we are really dealing with a system with an infinite number of degrees of freedom (for a free field, the infinite number of harmonic oscillators).

In weak coupling, transitions between these states will occur through tunneling. To estimate the tunneling amplitude, we first consider the case where the Higgs mass is of order the gauge boson mass (i.e., where the Higgs quartic coupling, λ, satisfies $\lambda \sim g^2$). In this case, we can rescale the various bosonic fields:

$$A_\mu \to \frac{1}{g} A_\mu \qquad\qquad \phi \to \frac{1}{g}\phi \;. \tag{11}$$

Then $\frac{1}{g^2}$ sits out front of the bosonic part of the lagrangian, i.e.,

$$\mathcal{L}_{bose} = \frac{1}{g^2}\left(-\frac{1}{4}F_{\mu\nu}^2 + |D\phi|^2 - V(\phi;\frac{\lambda}{g^2})\right) \;. \tag{12}$$

As a result, the classical equations don't involve g. We can obtain the tunneling amplitude by analogy to ordinary single particle quantum mechanics.[11] There, to compute the barrier penetration factor, one solves the equations of motion for the system with imaginary ("Euclidean") time, with boundary conditions such that in the far past the system is in one ground state, and in the far future it is in the other. In field theory, the corresponding solution is known as an "instanton"; for gauge theories, solutions of this type were first written down long ago.[12] We will denote such solutions generically by $\phi_{cl}(\vec{x},\tau)$, where ϕ denotes a generic bosonic field (i.e., it may refer to a scalar or to a gauge boson). Noting the form of the action in eqn. (12), the action of the instanton, S_{cl}, is necessarily proportional to $\frac{1}{g^2}$. As in ordinary quantum mechanics, the amplitude goes as $e^{-S_{cl}}$.

* This description is really only appropriate for weakly coupled theories. It is useful for discussing a theory like the electroweak theory, but it is not approprate for analyses of QCD. In addition, in order to form states with suitable clustering properties, it is necessary to superpose these so-called "n-vacua" to form Bloch waves, the "θ-vacua." All of the discussion of these lectures is easily rephrased in these terms.

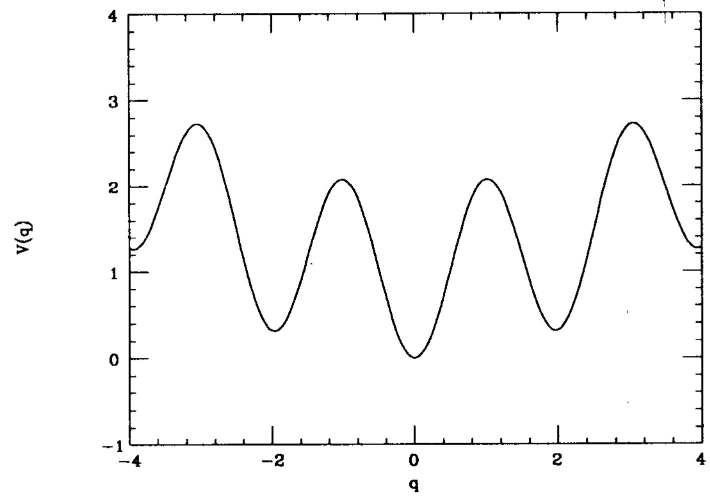


Figure 3. Schematic of potential in theory with fermions

By similar scaling arguments, we can determine the energy of the field configuration at the top of the barrier, the so-called "sphaleron." This is a static, unstable solution of the field equations, with a single negative mode. Again, since the coupling constant does not appear in the lagrangian, while the Hamiltonian is proportional to $\frac{1}{g^2}$, its energy is necessarily of order $\frac{M_W}{\alpha_2} \sim 10\; TeV$.

In theories with massless fermions, the situation is more complicated, due to the anomaly. Again, there are, at the classical level, an infinity of zero energy configurations, separated by a barrier, and labeled by an integer. However now, as a consequence of the anomaly, these different states carry different values of the non-conserved charge (baryon number). To see this, recall that the conserved charge is given by

$$\tilde{Q} = Q_5 + N_{CS} \;.$$

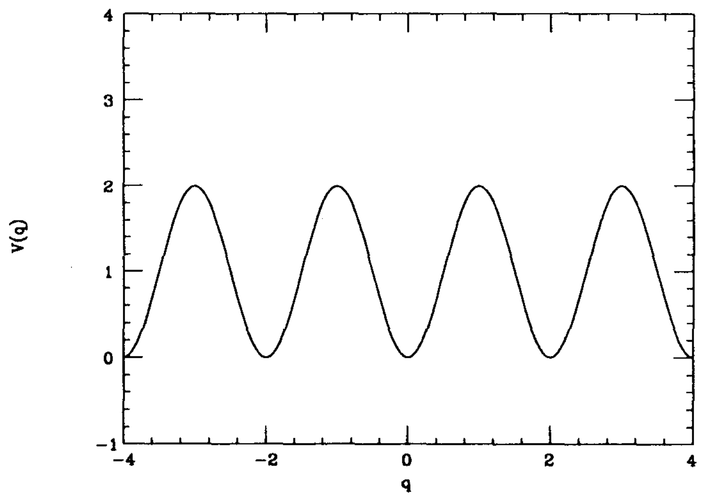


Figure 4. Particle in a periodic potential. V_0 is the potential barrier; ω is the curvature near the minima.

Thus bosonic field configurations which differ in their value of N_{CS} (those connected by the instanton solutions) differ also in their value of the charge Q_5, since they must have the same value of $\tilde{Q}$. There is no reason that states with different baryon number need be degenerate at the quantum level. Indeed, because of the exclusion principle, states with more baryons have higher energy. This is indicated in fig.3. The system can still tunnel between the various states. However, from the figure it is clear that this tunneling is accompanied by a change in the baryon number. This is the origin of the baryon-number violating effective interaction found by 't Hooft. This picture also makes clear the reason that the effect is so extremely small: the barrier penetration factor is proportional to the exponential of (minus) the instanton action, which is proportional to $\frac{1}{g^2}$.

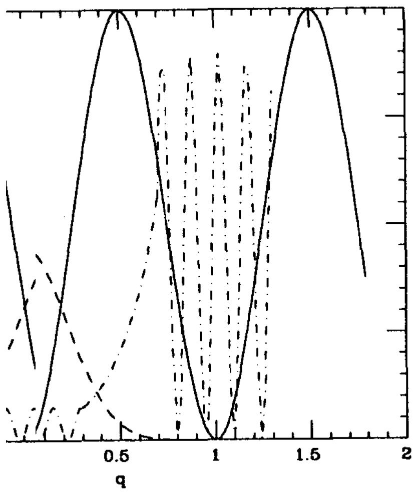
Since baryon number violation in the standard model is a problem of barrier penetration, the question naturally arises: while the effect is exponentially small at low energies and temperatures, might it be enhanced at energies or temperatures comparable to or greater than the barrier height? Clearly if one can kick the system in a suitable way, passage over the barrier will be a classically allowed process, and there will be no significant suppression. The question is: what constitutes a suitable kick? To get some insight into this problem, consider a one dimensional quantum mechanical system, with a periodic potential (fig. 4). At zero temperature, the system can pass from one well to another by tunneling. Suppose, now, that the system is placed in thermal equilibrium with a heat bath at a temperature $T \gg \omega_0$. This is the regime of classical statistical mechanics, so one can use the Boltzmann distribution. The probability to find the system near the top of the barrier is simply $e^{-\beta V_0}$. Thus the rate for passage over the barrier quickly becomes much larger than the tunneling amplitude for the low lying states. At temperatures greater than or of order the barrier height, the rate simply becomes of order one.

KRS suggested that this is a correct analogy for field theory.[4] In particular, once the temperature of the system is large compared to the typical masses, the barrier penetration rate, they argued, can be estimated using classical reasoning, and is proportional to $e^{-\beta E_{sp}}$. There has been much discomfort with this reasoning. In particular, it has been suggested that the simple quantum mechanics analogy is not relevant to field theories with their infinite number of degrees of freedom. However, in the last few years, the correctness of this picture has become firmly established.[5,6]

Can one similarly enhance the rate by scattering particles with very high energies? In particular, can scattering of two high energy particles lead to passage over the barrier and to baryon number violation? Before attempting to attack this question in field theory, let us again try to construct a quantum mechanical analog. A suitable model for this problem requires at least two degrees of freedom, one of high frequency and one of low frequency. Such a model has been developed and analyzed by Singleton, Susskind and Thorlacius.[13] We will consider a simpler model here: couple the original quantum mechanical variable, q, to a high frequency oscillator, with coupling

$$H_I(q) cos \omega t$$

with $\omega \gg \omega_0$. Assuming the coupling is weak, we can evaluate the cross section for transitions across the barrier using Fermi's golden rule. For this, we need the



ateshe quantum mechanics problem. Initial state welnal state is highly excited centered in next.

$f|H_I|i>$

,ed g. 5, for the initial state we take a harmonic ed ıne well, while for the final state we take a the t well. The initial state wave function looks

$$_i \sim e^{-\omega q^2} \quad . \tag{13}$$

hae form

$$e^{i\sqrt{2(E-V)}} \quad \sim \mathcal{T} e^{i\sqrt{2E}q} \quad . \tag{14}$$

tioctor. Because of the rapid oscillation of the cre exponentially with energy,

$$< _I|i > \sim \mathcal{T} e^{-\frac{E}{\omega}} \; . \tag{15}$$

The barrier penetration factor, $\mathcal{T}$, is given by

$$\mathcal{T} \sim e^{-\int dq \, \sqrt{2(V-E)}} \tag{16}$$

where the integral is taken between the turning points. At $E = 0$, $\mathcal{T}$ is exponentially small. As the energy increases, $\mathcal{T}$ grows exponentially with energy at first. However, it does not grow indefinitely. In particular, it can never become larger than one. Thus, from eqn. (15), we see that the rate for passage over the barrier is always exponentially small in this model. This seems a reasonable result: classically we know that it is difficult to excite one oscillator with another of very different frequency. The model of ref. 13 exhibits very similar features. In the high energy scattering problem, as we will see in more detail below, we need to couple high momentum (and energy) quanta to the relatively low momentum modes associated with passage over the barrier. Thus the analog model suggests that rates for such processes, while they may grow somewhat with energy, should never become large.

Instanton amplitudes, on the other hand, have been argued to give a quite different result.[7,8] In the leading WKB or semiclassical approximation, instanton amplitudes are in fact independent of energy. We can derive this result very easily, even without a detailed understanding of the instanton solutions themselves. Fermions are inessential for this problem; we can simplify things further by considering scalars only. Then the textbook LSZ analysis tells us that the cross section for scattering of n scalars of momenta $p_1, \dots, p_n$ is given by

$$S(p_1, \dots p_n) = (p_1^2 - m^2) \dots (p_n^2 - m^2) \int d^n x_1 d^n x_n e^{ip_1 \cdot x_1 \dots ip_n \cdot x_n} < \phi(x_1) \dots \phi(x_n) > . \tag{17}$$

Since Landau, we have also known that scattering amplitudes are analytic in the kinematic invariants, and that it is convenient to analytically continue to Euclidean momenta, $p_i^2 < 0$. From a space-time point of view, this is equivalent to continuing to imaginary time, $t \to -it$, or Euclidean space. It is precisely in Euclidean space that instantons enter. The Green's function in eqn. (17) can be written as a Feynman path integral,

$$\int [d\phi] e^{-\frac{S[\phi]}{h}} \phi(x_1) \dots \phi(x_n) \quad . \tag{18}$$

Here the integral is written over all possible histories of the field, $\phi(\vec{x}, t)$, where, because we are in Euclidean space, there is no factor of i in the exponential. In the classical limit ($h \to 0$), the path integral is dominated by the stationary points of the exponential, i.e., the field configurations, ϕ_{cl}, for which $\frac{\delta S}{\delta \phi_{cl}} = 0$; these are just the classical solutions of the Euclidean equations of motion, i.e., the "instantons." (More precisely, the instantons are the classical solutions of finite action.) Near these points in field space, we expand $\phi = \phi_{cl} + \delta\phi$. The integral over $\delta\phi$ is approximately gaussian.

The leading semiclassical approximation consists of keeping terms up to quadratic order in $\delta\phi$. Any finite action solution is necessarily localized in space and time, so instanton solutions are not translationally invariant. Translational invariance is reflected by the fact that if ϕ_{cl} is a solution, so is $\phi_{cl}(x - x_0)$. To obtain translationally invariant results for physical quantities, it is necessary to integrate over x_0. Thus in the leading semiclassical approximation, the Green's function appearing in the LSZ formula is given by

$$< \phi(x_1) \dots \phi(x_n) >= c e^{-S_0} \int dx_0 \phi_{cl}(x_1 - x_0) \dots \phi_{cl}(x_n - x_0) \; . \tag{19}$$

The constant c represents the result of performing the Gaussian integrals over the fluctuations $\delta\phi$. Actually, we need the Fourier transform,

$$c e^{-S_0} \int dx_1 \dots dx_n dx_0 \phi_{cl}(x_1 - x_0) \dots \phi_{cl}(x_n - x_0) e^{ip_1 \cdot x_1 \dots ip_n \cdot x_n} \quad . \tag{20}$$

Shifting $x_i \to x_i + x_0$, the amplitude factorizes, giving simply

$$c e^{-S_0} \int dx_0 e^{-i \sum p_i \cdot x_0} \int dx_1 \phi_{cl}(x_1) e^{ip_1 \cdot x_1} \dots \int dx_n \phi_{cl}(x_n) e^{ip_n \cdot x_n} \quad . \tag{21}$$

Because of this factorization, the amplitude is only a function of the various p_i^2, and is completely independent of the invariants $p_i \cdot p_j$ for $i \neq j$! As a result, the analytic continuation back to the mass shell, $p_i^2 = m^2$, is almost trivial. It is not hard to show that the Green's function has the correct poles. At large distances, the solution ϕ_{cl} must behave like the free Green's function, i.e., $\phi_{cl}(\vec{x} - \vec{x}_0) \to aG(\vec{x} - \vec{x}_0)$. The fourier transform thus has poles at $p^2 = m^2$. On the other hand, the amplitudes are independent of the center of mass energy, and all other interesting invariants! To obtain the cross section, one just needs to know how the amplitude for the emission of n particles depends on n, and the form of the n particle phase space. For the total cross section, one obtains[9]

$$\sigma = \sum \sigma_n \sim e^{-S_0} e^{\frac{E}{m_H}} \quad . \tag{22}$$

This grows rapidly with energy. If we take the result literally up to energies of order $\frac{E}{\alpha}$, the second factor can become comparable to the suppression, e^{-S_0} and the cross section will become large. In fact, the authors of ref. 9 have argued that this formula is not valid at such high energies, but simple improvements in the analysis give a result which grows even faster with energy! (A particularly compelling critique of this calculation is given in ref. 14.)

The analysis above did not include the fermions. In the standard model, as described earlier, the instanton represents a tunneling between nearly degenerate states of different baryon number. As a result, when fermions are included, the instanton amplitudes are non-vanishing only for Green's functions (S-matrix elements) which violate baryon number by a suitable amount. Moreover, the largest contribution to the total cross section in eqn. (22) comes from states with a large number ($\mathcal{O}(\frac{1}{\alpha_2})$) gauge or Higgs bosons. Thus the instanton analysis predicts a large rate for baryon number violating events with huge numbers of Higgs or W's and Z_0's in the final state. We have also ignored a number of other complications up to now. The fermions and gauge bosons give some modification of the energy dependence of the result, as do the various "collective coordinates" which we have suppressed (apart from the translational ones). When one takes these effects into account the basic picture remains the same.

The Euclidean calculation is in some ways rather unnerving. The continuation to Minkowski space is completely mysterious. For example, we have no sense what are the relevant configurations in real time. Moreover, it is not at all clear over what range of energies the analysis is valid.[15] Fortunately, we can give an alternative derivation of this result directly in Minkowski space. This derivation makes clear why the instanton calculation gives a growing cross section. Just as in the simple quantum mechanical example, there is a competition between the growth of a barrier penetration factor, and suppression due to the cost of coupling to a low frequency (momentum) mode. This calculation also makes clear that the rate is exponentially small at all energies.

The idea is to take the full field theory, with its infinite number of degrees of freedom, and truncate it to a single degree of freedom, i.e., to an ordinary quantum mechanics problem. For notational reasons only, we will idealize the problem by considering a theory with only scalars, and we will ignore most of the collective coordinates. Consider, then, a classical solution, $\phi_{cl}(\vec{x},\tau)$. This solution interpolates between two different vacua of the standard model. At $\tau=-\infty$, ϕ tends to one vacuum configuration; at $\tau=+\infty$, ϕ tends to the next. We can think of τ as parameterizing a set of field configurations which smoothly interpolate between one vacuum and the next. Introduce a coordinate, q, by letting $\tau=\tau(q)$, with $\tau(0)=-\infty$ and $\tau(1)=\infty$. The potential energy as a function of q looks as in fig. 2. Now we truncate the field theory by keeping *only* those field configurations parameterized by q. We treat q as an ordinary coordinate, i.e., we let $q=q(t)$. The lagrangian is then a function of q and the velocity, $\dot{q}$, given by

$$g^2\mathcal{L}(q,\dot{q})=\int d^3x\mathcal{L}(\partial_\mu\phi_{cl}(\vec{x},q),\phi_{cl}(\vec{x},q)) \ . \tag{23}$$

Provided that the dependence of τ on q is chosen appropriately, the lagrangian takes the form

$$g^2\mathcal{L}(q,\dot{q})=\frac{1}{2}\dot{q}^2-\frac{1}{2}m^2q^2-\mathcal{O}(q^4) \ . \tag{24}$$

To work out the connection between τ and q, recall that for $|\tau|\to\infty$, the classical solution behaves like the free propagator, i.e.,

$$\phi_{cl}\sim e^{-m\sqrt{\vec{x}^2+\tau^2}} \quad . \tag{25}$$

Thus

$$\int dx\dot{\phi}_{cl}^2\sim\dot{\tau}^2e^{-2m\tau} \tag{26}$$

so

$$q\sim e^{-m\tau} \quad . \tag{27}$$

This truncation of the field theory may seem very drastic, but we have not really lost anything. The truncated model contains *all* of the physics of the original instanton computations. The q system possesses an instanton solution in Euclidean time which describes the tunneling from one well to another. If we denote this solution by $q_{cl}(\tilde{\tau})$, then $\tau(q_{cl}(\tau))=\tau$. As a result, first replacing the field theory Green's functions appearing in the LSZ formula,

$$<\phi(x_1)\dots\phi(x_n)>$$

by

$$<\phi_{cl}(\vec{x}_1,q(\tau_1))\dots\phi_{cl}(\vec{x}_1,q(\tau_1))>$$

and then replacing q by q_{cl}, we obtain precisely the same quantity as in the leading semiclassical approximation in the field theory. Thus the scattering amplitudes computed in the leading semiclassical approximation to the truncated system are precisely the same as those calculated in the leading semiclassical approximation to the full field theory. In particular, they are independent of energy, and yield total cross sections which grow exponentially with energy.

On the other hand, since we are now dealing with a quantum mechanical system with only a single degree of freedom, it is easy to treat the problem directly in Minkowski space, and to understand the behavior found in the instanton computations. The two particle initial state of the scattering problem is

$$|i>=\lim_{T\to-\infty}\int dx_1dx_2\phi(x_1,T)\phi(x_2,T)|0>e^{ip_1\cdot x_1+ip_2\cdot x_2} \ . \tag{28}$$

Making our truncation, this becomes

$$|i> = \lim_{T\to-\infty}\int dx_1 dx_2 \phi_{cl}(x_1, q(T))\phi_{cl}(x_2, q(T))|0> e^{ip_1\cdot x_1 + ip_2\cdot x_2}$$

with a similar expression for the final state. These expressions are written in the Heisenberg picture. Switch to the Schrodinger picture, and take the overlap with the state of the q system of energy E. Because of the factor $e^{i(\omega_1+\omega_2)T}$, only the state with $E = \omega_1 + \omega_2$ gives a non-zero amplitude as $T \to \infty$. Thus we need to compute the overlap

$$\int dx_1 dx_2 < E|\phi_{cl}(x_1, q)\phi_{cl}(x_2, q)|0 > e^{i\vec{p}_1\cdot\vec{x}_1 + i\vec{p}_2\cdot\vec{x}_2} \quad . \qquad (29)$$

Even before attempting to evaluate this matrix element, it is clear that it is small for large $|\vec{p}|$. $\phi_{cl}(\vec{x}, q)$ is a smooth function of $\vec{x}$, with some characteristic scale, $\rho(q)$. Thus its fourier transform behaves as

$$\phi_{cl} \sim e^{-p\rho} \, .$$

ρ has a minimum as a function of q; in the standard model, roughly speaking, this is m_W^{-1}. As a result, it is difficult, at any energy, to couple high momentum modes to q.

How can we reconcile this with the instanton result? In fact, for energies $m \ll E \ll \frac{1}{g^2}m$, the truncated model yields *precisely* the instanton result. Recall that for small q (large τ)*

$$g^2 V(q) = \frac{1}{2}m^2 q^2 \ .$$

Thus requiring that $V \sim E$ gives $q^2 \sim \frac{g^2 E}{m^2}$. This means $\tau \sim m^{-1} ln(\frac{g^2 E}{m})$. On the other hand, for such τ, the classical solution has the form $e^{-m\sqrt{\vec{x}^2+\tau^2}}$. Its fourier transform is thus of order $e^{-E\tau} \sim (g^2E)^{\frac{E}{2m}}$. This can also be seen by noting that $e^{-E\tau} \sim q^{\frac{E}{m}}$. We need the matrix element of this operator between the ground state of the q system and the state of energy E. Since in this regime the potential is approximately harmonic, the state of energy E is just the n'th harmonic oscillator level, with $n = \frac{E}{m}$. Noting the form of the lagrangian in eqn. (24), we see that

$$< 0|q^n|n > \sim \sqrt{n!}g^n \sim (g^2E)^{\frac{E}{m}} \quad . \qquad (30)$$

This evaluation agrees with the intuition that the coupling to the initial state is suppressed because the sphaleron contains a large number of quanta. A similar factor arises for the final state. In addition to these overlap factors, however, it is also necessary to compute the barrier penetration factor. The WKB approximation gives

$$\mathcal{T} \sim e^{-\int dq\sqrt{2(V-E)}} \sim e^{-\int dq\sqrt{2V}} \times e^{E\int \frac{dq}{\sqrt{2V}}} \quad .$$

The leading behavior of the integral is obtained by approximating V by $\frac{1}{2}m^2q^2$, while cutting off the q integration at small q by E, and at large q by $\frac{1}{g^2}$. Then

$$\mathcal{T} = e^{-S_0}(g^2E)^{-\frac{E}{m}} \quad . \qquad (31)$$

Thus the scattering amplitude is indeed constant! This arises because the growth of the WKB factor initially precisely compensates for the difficulty of pumping energy into the q mode.

It is also very easy to reproduce the instanton result for the total cross section in the truncated model. From the point of view of the q system, the final state is simply the excited state of energy E. The field theoretic calculation of the amplitude to obtain a particular final state is obtained from this by taking a suitable overlap. Thus the total cross section is just the square of the amplitude that the original two particle state produces a final state of the q system in the other well. This amplitude is the product of the factor $\mathcal{T}$, and the factor from the overlap of the *initial* two-particle state with the state of energy E, i.e.,

$$\sigma_t \sim |\mathcal{T}|^2 (g^2E)^{-\frac{E}{m}}$$

$$\sim e^{-2S_0}e^{\frac{E}{m}} \quad . \qquad (32)$$

* It is not hard to show that the mass which appears below is precisely the meson mass.

This agrees precisely with the instanton result.† From this viewpoint, it is clear, however, that the answer is *never* large. The full amplitude is always the product of the barrier penetration factor, which is bounded by unity, and the small coupling to the q mode. The amplitude at all energies is bounded by $e^{-\frac{c}{g^2}}$, where c is some numerical constant. This may be much larger than the low energy result, but it is clear that the truncated system leads, at *all energies* to cross sections far too small to be seen in accelerator experiments.

One can object that the truncated model cannot be trusted at energies comparable to the barrier height. Indeed, there are a variety of ways in which the approximations involved can break down. For example, at very high energies, other trajectories, far from the original instanton path through configuration space, may be important, and give larger values for the cross section. On the other hand, ***the truncated model includes all of the physics which went into the original predictions of large cross sections. Thus, since this model in fact predicts only small cross sections, it is fair to say that no argument has been advanced that the cross section should be large at high energies!*** Moreover, the analysis of this system makes clear that no other trajectories in field space are likely to lead to large amplitudes either. The basic problem can be stated in a variety of ways. First, the barrier height associated with a particular configuration of scale ρ goes as $\frac{1}{\alpha\rho}$. To have any hope of enhancing the tunneling rate, the energy (and momenta) of the scattered particles must be at least this large. On the other hand, the coupling of a mode of scale ρ to a particle of momentum p goes as $e^{-p\rho} \sim e^{-\frac{1}{\alpha}}$ for such energies. Indeed, Susskind[16] has argued that at high energies the truncated model necessarily drastically overestimates the rate, once interference effects are taken into account.

Unfortunately, then, we must conclude that baryon number violation is not something one will see in accelerator experiments. Before closing, we should mention that another process which has been mentioned frequently[4] is similarly suppressed.[6] This is the decay of a very heavy fundamental particle by passage over the barrier with accompanying violation of baryon and lepton number. Consider, again, the truncated system. Now one has an initial state containing the heavy particle, with the q system in its ground state. The final state of interest involves a highly excited state of the q system in another well. As is clear from our simple quantum mechanics example, the overlap of the initial and final states appearing in this scattering amplitude will be exponentially small, i.e., the rate will go as $e^{-\frac{M}{m_W}}$, where M is the mass of the heavy particle. However, in strongly coupled theories where the heavy particle is composite, the decay may be rapid, as discussed in ref. 17. In particular, in technicolor theories, in the large N limit, one can show that technibaryon decay through the anomaly is unsuppressed. However, in the limit considered by these authors, the technibaryon is large (with size m_W^{-1} [17]) and the q degree of freedom is highly excited. Indeed, the production of this particle in, say, quark quark scattering is exponentially small due to the small form factor of this particle at momenta of order its mass–again it is hard to make excited states of the q system.

The lesson from all of this is that baryon number violation in the standard model is interesting, but its prime relevance is to the early universe. There, the fact that baryon violating interactions are in equilibrium until relatively late (temperatures of order 100's of GeV or so) means that any baryon and lepton number assymmetry created at very early times (e.g., in a grand unified theory or string theory) will be drastically altered from its initial value. This by itself is not terribly dramatic. If the underlying theory preserves $B - L$, then initially $B - L = 0$, and any baryon or lepton number created at this early epoch will be wiped out. However, most grand unified theories do not preserve $B - L$; even the simplest $SU(5)$ model only does so approximately, in the limit that non-renormalizable terms are not included. Thus one obtains an additional constraint on the parameter space of unified models. More interesting is the possibility that the baryon number violation of the standard model is itself responsible for the observed asymmetry between matter and antimatter. Various interesting suggestions along these lines have been made, but I do not believe this subject has been exhausted.[4,18–21] It is almost certainly necessary to modify the standard model so as to include more CP violation, but this can

† Agreement also holds when care is taken with subleading terms. Similar agreement is found when greater care is taken with the handling of the collective coordinates of the standard model problem, particularly the instanton scale size.

probably be accomodated. Recent calculations also suggest that the transition rate depends in a drastic way on the Higgs mass.[22] Thus baryon number violation in the standard model represents an area of standard or near standard model physics which may still yield surprising new physics.

Acknowledgements: I wish to thank my collaborators, T. Banks, G. Farrar, D. Karabali and B. Sakita, as well as L. Susskind, for enlightening discussions of these issues. This work supported in part by DOE contract DE-AC02-83ER40107.

REFERENCES

1. A.D. Sakharov, JETP Lett. 6 (1967).
2. G. 't Hooft, Phys Rev. Lett. **37** (1976); Phys. Rev. **D14** (1976) 3432.
3. N. Manton, Phys. Rev. **D28** (1983) 2019; F. Klinkhammer and N. Manton, Phys. Rev. **D30** (1984) 2212.
4. V.A. Kuzmin, V.A. Rubakov, and M.E. Shaposnikov, Phys. Lett. **155B** (1985) 36.
5. P. Arnold and L. McLerran, Phys. Rev. **D37** (1988) 1020; E. Mottola and A. Wipf, Phys. Rev. **D39** (1989) 588; S. Khlebnikov and M. Shaposnikov, Nucl. Phys. **B308** (885) 1988; J. Ambjorn, M. Laursen, and M. Shaposnikov, Phys. Lett. **B197** (1987) 491; A. Ringwald, Phys. Lett. **201B** (1988) 510; D. Grigoriev, V. Rubakov and M. Shaposnikov, Phys. Lett. **216B** (1989) 172.
6. M. Dine, O. Lechtenfeld, B. Sakita, W. Fischler and J. Polchinski, CCNY-HEP-89/18 (1990).
7. A. Ringwald, Nucl. Phys. **B330** (1990) 1.
8. O. Espinosa, Cal Tech preprint CALT-68-1586 (1990).
9. L. McLerran, A. Vainshtein and M. Voloshin, Phys. Rev. **D42** (1990) 171.
10. T. Banks, M. Dine, G. Farrar, D. Karabali and B. Sakita, Rutgers preprint RU-90-20 (1990).
11. See, for example, L. Schiff, *Quantum Mechanics*, McGraw Hill, New York (1968).
12. A.A. Belavin, A.M. Polyakov, A.S. Schvartz and Y.S. Tyupkin, Phys. Lett. **59B** (1975) 85.
13. R. Singleton, L. Susskind and L. Thorlacius, SLAC-PUB-5225 (1990).
14. K. Aoki and P. Mazur, UCLA preprint UCLA/89/TEP/67 (1989).
15. P. Arnold and M. Mattis, Los Alamos preprint LA-UR-90-1218 (1990); J. Cornwall, Phys. Lett. **B243** (1990) 271; A.H. Muller, Columbia preprint CU-TP-454 (1990); L. Yaffe, University of Washington preprint DOE-ER-40423-PT-90-07 (1990).
16. L. Susskind, private communication.
17. V.A. Matveev,V.A. Rubakov,A.N. Tavkhelidze and M.E. Shaposnikov, Sov. Phys. Usp. **31** (916) 1988.
18. M.E. Shaposnikov, Nucl. Phys. **B287** (1987) 757; Nucl. Phys. **B299** (1988) 797; A.I. Bochkarev, S. Yu. Khlebnikov and M.E. Shaposnikov, Nucl. Phys. **B329** (1990) 490.
19. L. McLerran, Phys. Rev. Lett. **62** (1989) 1075.
20. A. Cohen, D. Kaplan and A. Nelson, "Baryogenesis at the Weak Phase Transition," ITP preprint NSF-ITP-90-85 (1990).
21. N. Turok and J. Zadrozny, "Dynamical Generation of Baryons at the Electroweak Transition," Princeton preprint PUPT-90-1183.
22. L. Carson, X. Li, L. McLerran and R. Wang, "Exact Computation of the Small Fluctuation Determinant Around a Sphaleron," Minnesota preprint TPI-MINN-90/13T (1990).