Fermions near two-dimensional surfaces

Mark Burgess and Bjørn Jensen
Institute of Physics, University of Oslo, P.O. Box 1048, Blindern, 0316 Oslo 3, Norway
(Received 4 February 1993)

We review work on constrained systems in which (3+1)-dimensional field theories are reduced to effective (2+1)-dimensional ones. Known results are extended to encompass the Dirac equation and the nonrelativistic limit is examined. We discuss to what extent this system can really be made two dimensional and obtain a lower bound on the thickness. Some comments are made about recent theories involving fractional statistics.

PACS number(s): 03.65.−w

I. INTRODUCTION

Activity in low-dimensional physics has intensified in recent years, largely due to the discovery of the fractional quantum Hall effect and high temperature superconductivity, which are widely believed to be two-dimensional phenomena [1,2]. The starting point for descriptive theories is a (2+1)-dimensional system in which no residue of the third space-time direction survives. Work by Jensen and Koppe [3] and da Costa [4] has emphasized that a two dimensional system would in general have some knowledge of its surrounding three-dimensional space. In this paper we extend the work of Refs. [4–7] and comment upon its application to the more recent theories which have been used to study the quantum Hall effect and high-temperature superconductivity.

The literature on two-dimensional physics is not new. In semiconductor device technology, thin wafers may be thought of as confining electrons to an effectively two-dimensional space, trapped in three dimensions. See the review by Ando, Fowler, and Stern [8] for a considerable list of references. Since no physical system is truly two-dimensional, it is natural to ask how much such an ostensibly two-dimensional system would be able to "sense" of its three-dimensional embedding space. If the surface is curved, either instrinsically or extrinsically, then the system will at least be able to "feel" this curvature in the form of some mass or effective potential. This effect may be studied by expanding around the two-dimensional space for vanishingly small excursions in the third direction. The procedure has been studied in both the first quantized [3,4] and path-integral [6,7] formalisms for the Schrödinger equation. This is probably the case with the strongest validity in low-dimensional systems where there is little reason to think *a priori* that relativistic effects would be important. More recent work on chiral spin liquids and quantum Hall systems have used relativistic formulations, however (see [2] and references contained therein for a overview), and some authors have even considered anyons in the framework of Dirac theory [9]. Relativistic fermions also appear as the relevant low-energy degrees of freedom in the strong-coupling limit of the Hubbard model and strongly correlated systems [10]. Since a two-dimensional theory must be confined in some way by a rather large potential, it is important to consider the relativistic effects which might result. We have in mind the so-called "Klein's paradox" [11,12] which we shall refer to in Sec. III to obtain some limits on how two-dimensional a system can really be in the strict sense. This limitation has not previously been considered in the (2+1)-dimensional literature, although an application of the Dirac equation has been briefly treated in Ref. [13]. Other issues in Dirac theory are also of interest. In reducing the Dirac equation from 3+1 dimensions to 2+1 dimensions one may transform from the **4** representation to the fundamental **2** representation of the Clifford algebra, which is given by 2×2 matrices. In the process of reduction one normally expects to acquire two copies of the **2** matrices with opposite signs. This is required for invariance under parity transformations, since the angular momentum vector $\boldsymbol{\sigma}$ is a pseudovector. In the presence of a magnetic field, or some other mechanism under which P invariance is naturally broken, one might expect to obtain an asymmetric theory naturally expressible in the **2** representation. It is interesting to discover whether or not the type of two-dimensional surface a physical system is restricted to could induce a breach of parity invariance, leading to an asymmetric theory. If such a residue of non-P invariant Dirac fermions survives, then a Chern-Simons Lagrangian could arguably be induced by radiative corrections in a field theory [14]. This is of particular interest to recent work on the theory of anyons and fractional statistics [15–17].

II. GEOMETRY

Our aim is to probe three-dimensional space from the viewpoint of a two-dimensional system. We begin by defining the geometry of an arbitrary surface S in $\mathbb{R}^3$ (see Fig. 1 in Ref. [5]). Since our aim is to end up inside this surface it is useful to expand a set of vectors $\mathbf{X}$ which live in the immediate neighborhood outside the surface S to a set of vectors $\mathbf{x}$ which define the surface itself. We can use these as the basis for two sets of three-dimensional coordinates X^i and x^i, which we shall em-

1050-2947/93/48(3)/1861(8)/$06.00

ploy in the analysis. X^i is essentially equal to x^i plus a small perturbation in the direction normal to S. We shall use small letters to refer to geometrical objects in these "small-x" coordinates.

Let $\hat{\mathbf{e}}_\mu$ ($\mu = 0, \ldots, 3$) be a basis for our surface coordinates x, so that a vector in x has components x^i, x^3. The basis vector $\hat{\mathbf{e}}_3 \equiv \hat{\mathbf{n}}$ is normal to the surface. The space-time line element may be written in either of our two coordinate regimes [4,5]

$$ds^2 = -dt^2 + G_{ij}\, dX^i\, dX^j + (dX^3)^2 \tag{1}$$

$$= -dt^2 + g_{ij}\, dx^i\, dx^j + (dx^3)^2, \tag{2}$$

$i, j = 1, 2$. Our aim is now to relate these two metrics by expressing G_{ij} as a perturbation around g_{ij}. If $\nabla_{\hat{\mathbf{e}}_i}$ is a covariant derivative in the given basis $\hat{\mathbf{e}}_\mu$ then we may define the extrinsic curvature of our surface as follows. Since the derivatives of the normals lie in the tangent space we have

$$\nabla_{\hat{\mathbf{e}}_i} \hat{\mathbf{e}}_3 = -K_i(\hat{e}_j) = -K_i{}^j \hat{\mathbf{e}}_j \ . \tag{3}$$

$K_i{}^j$ may also be expressed in terms of the three-components of an affine connection as in Ref. [6]. Moreover, a vector in X expresses the same information as a vector in x before we squeeze $x^3 \to 0$; thus we may write

$$\mathbf{X}(x^1, x^2, x^3) = \mathbf{x}(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2) + x^3\, \hat{\mathbf{e}}_3. \tag{4}$$

Thus the metric G_{ij}, which we shall assume to be diagonalizable, may be written in terms of the two-bein

$$G_{ij} = \mathbf{E}_i \cdot \mathbf{E}_j \ , \tag{5}$$

where

$$\mathbf{E}_i = \nabla_{\hat{\mathbf{e}}_i}(\mathbf{x}(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2) + x^3\, \hat{\mathbf{e}}_3) \tag{6}$$

$$= (\delta_i{}^j - x^3\, K_i{}^j)\hat{\mathbf{e}}_j. \tag{7}$$

Substituting this result and using the orthonormality of the basis vectors $(\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j) = g_{ij}$ one obtains that

$$G_{ij} = g_{ij} - 2K_{ij}x^3 + K_i^k g_{km} K_j^m (x^3)^2. \tag{8}$$

The determinant of this quantity will be useful later. To order $(x^3)^2$ we have

$$G = g[1 - 2\,\mathrm{Tr}K x^3 + (2K + \mathrm{Tr}K^2)(x^3)^2 + \cdots] \ , \tag{9}$$

where $G = \det(G_{ij})$, $g = \det(g_{ij})$ $K = \det K$, and $\mathrm{Tr}K = K_i^i$. The square root of this determinant is also required for the invariant volume element. Expanding the binomial to order $(x^3)^2$ we have

$$G^{1/2} \equiv g^{1/2}\, \xi \ , \tag{10}$$

where

$$\xi = [1 - \mathrm{Tr}K\, x^3 + K(x^3)^2 + \cdots]. \tag{11}$$

III. SCHRÖDINGER EQUATION

The nonrelativistic treatment of this perturbative scheme has been considered by da Costa [4,5] in the first quantization and in the path-integral formulation very recently by Matsutani [6,7]. We begin by briefly recounting the work of these authors.

The action for the Schrödinger field in the coordinates x of our surface is

$$S = \int dV \left\{ \psi^\dagger\, i\partial_0\, \psi - \frac{1}{2m}(\nabla_k\psi)^\dagger(\nabla_k\psi) \right\} \ , \tag{12}$$

where $dV = dt\, d^3x\sqrt{g}$. Integrating the action by parts we obtain

$$S = \int dV \left\{ \psi^\dagger \left(i\partial_0 + \frac{1}{2m}\nabla^2 \right) \psi \right\} + \int dV\ \nabla^k(\psi^\dagger \nabla_k \psi) \ , \tag{13}$$

where the latter term arises as a result of our boundary S. The variation of this action leads to the field equation subjected to a boundary condition which is stipulated by the vanishing of the second term

$$\Delta(\psi^\dagger \nabla_3 \psi) \ , \tag{14}$$

where $\hat{e}_3$ is the normal to the surface S. This is satisfied provided the three-dimensional field is smooth as it passes through S.

The conserved current associated with the Schrödinger field implies that the probability density of finding a particle at x is proportional to $\psi^\dagger(x)\psi(x)$, so that the total number of particles (which is a constant) is given by

$$\int d^3x\sqrt{g}\ \psi^\dagger(x)\psi(x) = N. \tag{15}$$

One demands, under a change of coordinates, that this scalar quantity be preserved. Passing to coordinates X^i, we have

$$\int d^3X\sqrt{G}\ \phi^\dagger(X)\phi(X) = N. \tag{16}$$

From Eq. (10) it follows that

$$\phi(X) = \psi(X)\xi^{-1/2}. \tag{17}$$

We are now in a position to see what effect an excursion away from the surface has on the action. Reformulating (12) in X coordinates,

$$S = \int dV' \left\{ \phi^\dagger\, i\partial_0\, \phi - \frac{1}{2m}(\nabla_k\phi)^\dagger(\nabla_k\phi) \right\}, \tag{18}$$

where $dV' = dt\, d^3X\sqrt{G}$ so that

$$S = \int dV \xi \left\{ (\psi^\dagger \xi^{-1/2}) \left[i\partial_0 + \frac{1}{2m}\nabla^2 \right] (\psi\xi^{-1/2}) \right\} + \int dV'\ \nabla^3[(\psi^\dagger\xi^{-1/2})\nabla_3(\psi\xi^{-\frac{1}{2}})], \tag{19}$$

where $\nabla^2 = G^{-1/2}\partial_i\, G^{ij} G^{1/2}\partial_j$. The new boundary condition requires the continuity of K through the surface. This will be satisfied as long as the surface is not discontinuous. Pulling ξ through the derivatives introduces extra additive terms into the action which depend upon all of the coordinates $X^1, \ldots, X^3$. In the limit of a (2+1)-

dimensional theory we are interested in the case $x^3 \to 0$; thus we expand in terms of the x coordinates keeping terms which are zeroth order in x^3. [Note that since second-order derivatives are present we must know ξ to order $(x^3)^2$.] Thus, to zeroth order (in the limit $x^3 \to 0$) one finds that the following extra terms appear in the Lagrangian density:

$$\Delta\mathcal{L} = \lim_{x^3\to 0} \frac{1}{2m}\psi^\dagger \left[\frac{1}{4}\xi^{-2}(\partial_3\xi)^2 - \frac{1}{2}\xi^{-1}(\partial_3^2\xi)\right]\psi$$
$$= \mu_{\text{eff}}\psi^\dagger\psi \ , \tag{20}$$

$$\mu_{\text{eff}} = \frac{1}{2m}(\tfrac{1}{4}\text{Tr}K^2 - K). \tag{21}$$

The new term appearing as the "effective action" for the theory is a direct result of second-order derivatives probing the space around the surface. These terms survive even in the (2+1)-dimensional limit in the form of a minimal coupling to an effective potential. This potential makes a negative contribution to the energy. Matsutani interprets this as a chemical potential while da Costa treats it as an ordinary attractive potential which can lead to bound states. We prefer the latter interpretation since a system with nonconstant chemical potential is not in thermodynamic equilibrium and will therefore adjust itself by some transport of charge or matter until it is in equilibrium. Since a system cannot become uncurved simply by a flow of charge, it seems more natural to think of the potential only as a contribution to the energy.

We discuss now the issue of how the system is to be constrained to the surface S. We shall return to this more analytically in Sec. IV. Matsutani has addressed this issue in some detail by introducing a harmonic potential, effectively putting particles on springs and increasing the spring constant without limit. One begins by writing the wave function

$$\psi = \psi_{\|}(t, x^1, x^2)\psi_\perp(x^3). \tag{22}$$

The part of the wave function which represents behavior perpendicular to the two-dimensional surface is a function of x^3 only and is subject to a confining potential. The aim is to increase the strength of this potential so that the eigenfunction of the resulting motion become a delta function $\delta(x^3)$. To show this, one expands $\psi_\perp$ in eigenfunctions, parallel and normal to the two-dimensional surface

$$\psi_\perp = \sum_n a_n\,\psi_{\perp n}. \tag{23}$$

These functions have eigenvalues λ_n, which satisfy an equation of the form

$$\left[-\frac{1}{2m}\nabla_3^2 + \frac{1}{2}m\omega^2(x^3)^2\right]\psi_{\perp n} = \lambda_n\,\psi_{\perp n}. \tag{24}$$

As the strength of the confining potential ω increases, the states are confined more and more to the lowest-energy level, namely the ground state, which is characterized by the Gaussian wave function [6]

$$\psi_{\perp 0} \sim a_0\,\sqrt{m\omega/\pi}e^{-m\omega(x^3)^2}. \tag{25}$$

In the limit $\omega \to \infty$, the Gaussian function becomes the δ-function. One may therefore integrate over x^3 setting $x^3 = 0$ everywhere.

This semiclassical idea provides a formal procedure for the squeezing, but *physically* it is only an approximation. It is not possible to increase the strength of an arbitrary potential indefinitely without violating some physical condition on the system. We shall discuss this remark more carefully below in the relativistic treatment of the problem and attempt to determine some limits on the extent to systems may be regarded as two dimensional.

We note briefly that the form of the above eigenfunctions admits another confinement parameter, namely the mass. If the effective mass of particles were to increase as the particle moved away from the plane, i.e., $m = m(x^3)$, a similar confinement might be obtained. Whether or not this is a feasible picture depends partly on the nature of the model in question. If we imagine that our two-dimensional theory is really part of some layered system, with many similar systems on top of one another, then this may be given some substance (though note our remark in Sec. IV). In high-temperature superconductor models, for instance, there are many layers of different types, some of which are doped and therefore have a surplus of charge which could repel electrons from neighboring planes. If it is assumed that there is electromagnetic coupling between the layers, then electrons in S will behave like polarons, inducing a polarization cloud in neighboring planes, gaining an effective mass. This effect will give rise to a slight attractive or repulsive potential depending on the geometry, encouraging the particles to either traverse the gap between parallel layers or be repelled. However, if the layers are charged with the same sign as the normal charge carriers of the system, this effect will be small compared to the normal Coulomb repulsion at some distance. Approximating the field between neighboring layers to that of a parallel plate capacitor would give an effective mass for the polaron which would increase linearly with x^3 as it approached the neighboring plane. In the nonrelativistic treatment this is a reasonable point of view because the role of the mass and the coupling ω are the same. This is not the case in relativistic theory.

IV. DIRAC EQUATION

We now turn to the relativistic Dirac equation. The treatment here is complicated by the presence of the spin connection Γ_μ. On an arbitrary curved space-time manifold the Dirac equation is given by

$$(\gamma^a\nabla_a + m)\psi = 0. \tag{26}$$

The γ matrices are defined relative to a given tetrad field e^a_μ ($\omega^a = e^a_\mu dx^\mu$) and obey the usual anticommutation relation

$$\{\gamma^a, \gamma^b\} = 2\eta^{ab}. \tag{27}$$

Including a minimal coupling to a vector field A_μ the covariant derivative ∇ is defined by

$$\nabla_a = e_a{}^\mu(\partial_\mu - ieA_\mu - \Gamma_\mu) \ , \tag{28}$$

where the spin connection is

$$\Gamma_\mu = -\frac{1}{4}\gamma^a\gamma^b e_a{}^\nu(\partial_\mu e_{b\nu} - e_b{}^\lambda\Gamma_{\lambda\nu\mu}) \tag{29}$$

and $e_{b\nu} = G_{\nu\mu}e_b{}^\mu$. When $e_a{}^\mu\Gamma_\mu = 0$ we define $D_a = \nabla_a$. We may compute the Dirac equation in our coordinate system X relative to an orthonormal tetrade frame so that

$$\omega^0 = dt, \tag{30}$$

$$\omega^1 = G_{11}^{1/2}dx^1, \tag{31}$$

$$\omega^2 = G_{22}^{1/2}dx^2, \tag{32}$$

$$\omega^3 = dx^3. \tag{33}$$

The beins (tetrads) may therefore be read off from a knowledge of the metric.

Defining $\overline{\psi} \equiv \psi^\dagger\gamma^0$, the current $j^a \equiv \overline{\psi}\gamma^a\psi$ is covariantly conserved $\nabla_a j^a = 0$. The conserved current gives the probability density for the Dirac equation as $\overline{\psi}\gamma^0\psi = \psi^\dagger\psi$. Hence under a change of coordinates from x to X, probability conservation requires that the Dirac spinor $\psi(X) \to \psi(X)\xi^{-1/2}$ as for the Schrödinger field. We assume for simplicity that the surface has vanishing intrinsic curvature. We may then parametrize the surface by Cartesian coordinates. It is then easy to show that

$$\Gamma_3 = 0, \tag{34}$$

$$\Gamma_i = \frac{i}{2}\varepsilon_{ij}K_{ii}\sigma_j \cdot I \ , \tag{35}$$

where $i = 1, 2$ is summed and I is the identity for 4 spinor indices. We are now in a position to construct the covariant derivative explicitly. Since $\Gamma_3 = 0$ one sees that this is unnecessary however: there will be no contribution from the spin connection to the geometrical terms. We may therefore proceed to separate the Dirac equation into parallel and perpendicular parts. We have

$$(\gamma^A D_A + m)(\psi(X)\xi^{-1/2}) = 0 \ , \tag{36}$$

where $A = 0, \ldots, 3$. Pulling ξ through the derivatives we have

$$\left(\gamma^A D_A + \frac{1}{2}\gamma^3 \mathrm{Tr}K + m + O(x^3)^3\right)\psi(X) = 0. \tag{37}$$

We shall assume that the equation is separable, as in the non-relativistic case, and write

$$\psi(x) = \phi_\perp(x^3)\phi_\parallel(x^1, x^2, t)\hat{\psi} \ , \tag{38}$$

where $\hat{\psi}$ is a constant, four-component spinor. For later convenience we shall also define $\psi_\perp = \phi_\perp\hat{\psi}$, $\psi_\parallel = \phi_\parallel\hat{\psi}$, and $A^0 = A^0_\perp(x^3) + A^0_\parallel(x^1, x^2, t)$ and $m = m + \Delta m(x^3)$.

From (38)

$$\frac{(\gamma^0 D_0^\parallel + \gamma^i D_i + \frac{1}{2}\gamma^3 \mathrm{Tr}K + m)}{\phi_\parallel}\phi_\parallel\hat{\psi}$$

$$= -\frac{\gamma^0(-ieA_0^\perp) + \gamma^3\partial_3 + \Delta m(x^3) + \cdots O(x^3)}{\phi_\perp}\phi_\perp\hat{\psi}. \tag{39}$$

The usual argument for the separation of these equations is that the left-hand side is only a function of x^1, x^2, t while the right-hand side is only a function of x^3; thus both sides must be equal to some separation constant k. This is not quite correct, as we noted before, owing to the terms $\cdots O(x^3)^3$ which contain the curvature K_{ij}, which is a function of x^1, x^2. However, we are assuming that the excursions into x^3 are small owing to the physics of the problem; thus we shall for the time being adopt the standard procedure and neglect these higher-order terms. (We shall return to this point in Sec. VI, since it is not entirely trivial.) To eliminate x^3 from the problem we may identify $A_0^\perp$ with a squeezing potential $A_0^\perp = \frac{\lambda}{e}x^3$, so that nontrivial solutions, given by the vanishing of the determinant of the operator on the right-hand side of (40), are characterized by

$$\phi_\perp(x^3) \propto \exp\left[-\left(\frac{\lambda}{2}(x^3)^2 + kx^3\right) - \int dx^3 \ m(x^3)\right]. \tag{40}$$

The first term in the exponential falls off like a Gaussian function, i.e., a δ function in the limit of infinite λ as in Matsutani's scheme. The presence of a nonzero separation constant spoils this falloff by introducing a straightforward exponential decay with characteristic length k^{-1}. The mass parameter $m(x^3)$, which is initially at least linear in x^3, will also give rise to a Gaussian falloff. Thus increasing the coupling of this parameter will also have the desired effect of squeezing the system. The residual Dirac equation is

$$\left(\gamma^a D_a + \frac{1}{2}\gamma^3 \mathrm{Tr}K + (m - k)\right)\psi_\parallel = 0. \tag{41}$$

At this juncture we raise an issue associated with relativistic fields, which occurs when the confining potential becomes strong $\lambda \gg m^2$. As a strong potential is increased arbitrarily, it will cause the creation of particle pairs [11]. This phenomenon is connected to what is known as Klein's paradox in the literature [18]. It does not occur when the confinement is affected by $m(x^3)$, i.e., the potential transforms like the mass, since this only increases the energy gap in the rest frame of the particles. But for any four-vector-type potential this will be a relevant concern. The creation of particle antiparticle pairs happens essentially because there is enough energy available in the field to promote virtual particles into real ones. In the case of scalar particles, if the confining potential is nonhomogeneous and real particles are present, then this production may even be stimulated to increase [18,19]. This is a phenomenon which resembles the well-known laser effect [11]. In the literature it is known as superradiance [18], though it does not occur for fermions [18] essentially because the Pauli principle prohibits the occupation of the out-going states by more than one particle [20].

This production of particle-antiparticle pairs from the associated particle vacuum for sufficiently strong external fields [21] spoils the idea of absolute confinement in a relativistic system. One does not really escape the problem by only treating Schrödinger fields, which are clearly only a low-energy approximation to either a Dirac or Klein-Gordon field. One might argue that in very strong potentials particle fields should be treated relativistically precisely because of this limitation. On the other hand, since real layered materials clearly do not spontaneously expel particles, one should ask whether or not a potential which does not lead to such difficulties is capable of making a physical system truly two dimensional. The relativistic case therefore provides us with a natural limitation by which to gauge how two-dimensional a "two-dimensional" system can be. Let us examine this problem.

If the confining potential $\lambda \sim m^2$, then we are in danger of producing particle pairs. Although we are only after an order of magnitude estimate, we can be a little more precise. Assuming that the distribution around the surface S is Gaussian, one may define the probability P of all the particles being within $-d < x^3 < d$ by

$$P^{1/2} = \frac{\int_{-d}^{d} e^{-\lambda x^2}\, dx}{\int_{-\infty}^{\infty} e^{-\lambda x^2}\, dx}. \tag{42}$$

The integrals define the distance $d(\lambda)$ or conversely $\lambda(d)$. Since we have a bound $d\lambda < m$ this should specify the condition uniquely

$$\lambda(d) = \frac{1}{d^2} |\ln(1-P)|. \tag{43}$$

Reinstating SI units we obtain

$$\lambda = \frac{\hbar c}{d^2} |\ln(1-P)|, \tag{44}$$

$$d > \frac{\hbar c}{mc^2} |\ln(1-P)|. \tag{45}$$

$\frac{\hbar c}{mc^2}$ is notably the Compton wavelength of the particles. Substituting for the electron mass we have $\frac{\hbar c}{mc^2} \sim 10^{-13}$ m, which is clearly much smaller than the interatomic spacing $d_{\rm at} \sim 10^{-10}$ m.

A qualitative feel for the kind of charge density σ required to confine a gas of charged particles to two dimensions can be obtained by treating the system as a parallel plate capacitor. Here the electric field is given by $E = \sigma/\epsilon_0$ and thus we may write

$$d \sim \sqrt{\frac{\hbar c \epsilon_0}{e\sigma} |\ln(1-P)|}. \tag{46}$$

If we make the distance $d(\lambda)$ to be an order of magnitude less than the atomic spacing and require that *all* of the wave function be inside this distance [take $|\ln(1-P)| \sim 100$, for instance], then the necessary confining field is of the order of 10^{17} N/C. Compared to a typical atomic field $Ze/4\pi\epsilon_0 d_{\rm at}^2 \sim 10^{13} Z$ V/m this is enormous. It corresponds in fact to a charge density of 10^6 C m^{-2}, which is about 10^5 electron charges per square angstrom. We can start again with a more sensible estimate for the charge density. With $\sigma \sim 16$ C m^{-2}, that is, about 1 electron per square atom and $P = 0.8$ (47) gives $d \sim 4 \times 10^{-10}$ m. This is within the realms of credibility. The characteristic thickness of thin films and quantum wires is of the order of 100 Å [22], whereas the layer separation in high-temperature superconductors is of atomic proportions. In the latter case one must expect that the dominant forces involved in enforcing two dimensionality are due to chemical bonds which have a greater affinity for substances in the same surface than for substances in neighboring planes. At a distance of one Bohr radius from a typical atom, the electric field is well within the limits we have found above. On the other hand, a two-dimensional gas does not seem as plausible in the strict sense.

In Ref. [22] the authors obtain another limit of interest by determining an *upper* bound on d. This bound is fixed by the requirement of having at least one bound state arising from the induced potential in the nonrelativistic case. Combined with the above condition one has now a range of permissible distances over which bound states may be obtainable. As these authors remark, these distances are achievable with current technology.

V. 2 REPRESENTATION FOR SPINORS

In purely (2+1)-dimensional theories it is natural to express the Dirac equation in terms of the fundamental representation of the Clifford algebra. This may be constructed from the 2×2 Pauli matrices, which we shall denote the **2** representation. In 3+1 dimensions the fundamental representation is in terms of 4×4 (**4**) matrices. In 2+1 dimensions a **4** representation is also possible. Indeed the reduction of the Dirac equation in the preceding section must clearly lead to a 4×4 matrix expression. This **4** representation is therefore reducible in 2+1 dimensions. It is worth pointing out that the Pauli spin matrices are a pseudovector in 2+1 dimensions, which means that the Dirac equation expressed in terms of these matrices is not parity invariant in a massive theory. The two equations

$$(\not{D} \pm m)\psi = 0 \tag{47}$$

are inequivalent [14]. This problem does not arise in the **4** representation since the **4** representation reduces to two copies of the **2** representation symmetrized with respect to the mass. In an arbitrary curved surface it is not inconceivable that the curvature itself might break the reflection symmetry of the system in 2+1 dimensions, leading to an effective two-dimensional theory with a definite parity. We have in mind screw dislocations, for example, where one has torsion which behaves in many ways like a magnetic field. This issue is of some interest in connection with Chern-Simons theories of anyons, since it implies that a Chern-Simons term would be induced in the effective action at one-loop. This could be manifested by an asymmetry in the effective mass of the fermions after diagonalization. We begin by showing that there is, however, no violation of parity at least at the classical level if the Dirac equation can be diagonalized. The standard representation of the γ matrices in 3+1 dimensions is

$$\gamma^0 = i\begin{pmatrix} \underline{\mathbf{1}} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & -\underline{\mathbf{1}} \end{pmatrix}, \quad \gamma^i = -i\begin{pmatrix} \underline{\mathbf{0}} & -\underline{\sigma}^{\mathbf{i}} \\ \underline{\sigma}^{\mathbf{i}} & \underline{\mathbf{0}} \end{pmatrix}. \tag{48}$$

These satisfy a number of relations which are summarized in the appendix. To create a **4** representation in 2+1 dimensions it is sufficient to drop γ^3. The remaining three matrices satisfy the same relations, except that $\gamma_5 = -\gamma^0\gamma^1\gamma^2$ now commutes with the other matrices, as does its **2** representation. Consider the (3+1)-dimensional equation

$$(\gamma^{\hat{\nu}} D_{\hat{\nu}} + m + \mu\gamma^3)_{\mathbf{4}}\psi = 0 , \tag{49}$$

where $\hat{\nu} = 0, \ldots, 3$, D is an arbitrary covariant derivative, and μ is, in general, a real function of x^1, x^2 (in our case $\mu \propto \mathrm{Tr}K$). Writing this equation in terms of 2×2 matrices

$$\begin{pmatrix} iD_0 + m & i(\sigma^k D_k + \sigma^3\mu) \\ -i(\sigma^k D_k + \sigma^3\mu) & -iD_0 + m \end{pmatrix}\psi = 0. \tag{50}$$

This equation may be diagonalized by solving the secular equation for the eigenvalues. Formally,

$$\lambda = m \pm \sqrt{-D_0^2 + D_k^2 + \overline{\mu}^2 + \tfrac{1}{2}[\sigma^i, \sigma^j]D_i D_j} , \tag{51}$$

where $\overline{\mu}^2 = \mu^2 + \epsilon_{kj}\gamma_j(\partial_k\mu)$. The operator inside the square root has the form of a Klein-Gordon operator, the square root of which may be expressed in the form $\gamma^\mu D_\mu + A$, where the γ matrices are now in the **2** representation; thus squaring up and comparing

$$-D_0^2 + D_k^2 + 2A\gamma^\mu D_\mu + \gamma^\mu(\partial_\mu A) + A^2$$

$$= -D_0^2 + \partial_k^2 + \overline{\mu}^2. \tag{52}$$

If, additionally, we assume that the eigenvalues λ may be expressed in the form $\gamma^\mu\partial_\mu + M = 0$, then we have simultaneously $A+m = M$ and $-2AM + \gamma^\mu(\partial_\mu A) + A^2 = \overline{\mu}^2$, which implies

$$\gamma^i\partial_i M + M^2 = m^2 + \mu^2 + \epsilon_{kj}\gamma^j(\partial_k\mu). \tag{53}$$

The general solution to this equation is in general rather complicated and may be expressed in the form

$$M = \sum_n \frac{a_n}{(x - x_0)^n} + m + \sum_m b_m(x - x_0)^m. \tag{54}$$

A special solution may be obtained by noting that, since γ^i and $i = 1, 2$ are linearly independent matrices, we have three equations

$$M = \pm\sqrt{m^2 - \mu^2}, \tag{55}$$

$$\partial_i M = \varepsilon_{ij}\partial_j\mu. \tag{56}$$

For constant μ a related result has been derived in [23]. Noting that these equations are operator identities we have

$$\nabla^2 M + (\partial_i M)\partial_i = \epsilon_{ij}(\partial_i\mu)\partial_j + [\partial_1, \partial_2]\mu. \tag{57}$$

Using (57) this becomes

$$\nabla^2 M = J , \tag{58}$$

where $J = [\partial_1, \partial_2]\mu$. The Dirac equation is thus diagonalizable for any surface satisfying (58). We note, in particular, that the geometrical terms μ have the form of an effective mass or rest energy here (the effective mass is *reduced*), not an effective potential either in the form of a chemical potential or a four-vector potential. Moreover, the eigenvalues are symmetrical with respect to the sign of the mass and therefore there is no loss of parity invariance when diagonalization is possible, provided there is no impending physical reason for neglecting half of the solutions. (This may be the case in a strong magnetic field at low energies, where one may isolate spin-up and spin-down states.) In most situations $J = 0$, but this is not necessarily the case if there are conical singularities or holes in the manifold. Physically, this might occur around a vortex or disclocation of some kind. In this situation it is known that Aharanov-Bohm-type phases can result [15,24,16]. If the hole is not penetrated by any external flux, then the dynamics of the field determine the phase around the singularity [24]. Periodic bosons and antiperiodic fermions minimize the free energy [24,25]. This bears some qualitative resemblance to the phenomenon of fractional statistics [15,16,25].

In the nonrelativistic model considered in [4,5,22] the authors find a bound state in the so-called "bookcover" surface. This two-dimensional surface consists of a plane wrapped partially around a cylinder. It has no intrinsic curvature, but $\mathrm{Tr}K \propto 1/a$, $K = 0$. A bound state also seems to be possible in the relativistic theory insofar as the self-energy is reduced by the curvature of the manifold. This in itself seems counterintuitive: one might expect that particles in curved regions would be eager to escape their confinement, not be attracted by it. It is not difficult to see that the paradox is resolved when one takes account of whatever forces are actually reponsible for the curvature. Formally it is no more surprising than the lowering of potential energy in a gravitational well. What is curious is that the relativistic fermion system only probes the surrounding space to first order in the derivatives, whereas both Schrödinger fermions and bosons probe to second order. The nonrelativistic system seems to extract more information from the manifold than the relativistic one. In particular the determinant K, which appears nonrelativistically, is related through the Gauss-Bonnet theorem to the topological class of the manifold. (This is not as deep as it sounds, since it is the integral which decides the topology and not the local form of the curvature.) In the nonrelativistic limit, this information is returned as higher derivatives of the wave function and spin-connection appear.

VI. THE NONRELATIVISTIC LIMIT

In this section we remark on the nonrelativistic limit of the Dirac equation in the operator formalism. It is convenient to use the diagonalized **2** form for this purpose. As pointed out in [23] the two reduced blocks, after diagonalization, correspond to spin up and spin down when the positive- and negative-energy solutions are isolated. It is sufficient for our purpose to consider one of these blocks:

$$\begin{pmatrix} -i\partial_0 + M & -(D_1 - iD_2) \\ -(D_1 + iD_2) & i\partial_0 + M \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 0. \tag{59}$$

In Fourier space this becomes two positive-energy equations

$$(-E + M)u_1 - i\pi^* u_2 = 0, \tag{60}$$

$$-i\pi u_1 + (E + M)u_2 = 0, \tag{61}$$

where $\pi = p - eA - \Gamma$ is the Fourier transformed momentum. In this positive energy regime, the u_2 solutions are small by a factor v/c; eliminating these and writing $E = E_{\rm NR} + m$, where $E_{\rm NR}$ is the nonrelativistic energy, gives

$$-E_{\rm NR} - m + \sqrt{m^2 - \mu^2} + \frac{p^2}{2m} \pm \frac{eB}{2m} \pm T, \tag{62}$$

where T represents terms which can arise when there is torsion in the connection. Since, by our previous assumptions, $\mu \sim p \ll m$ and $M \sim m - \mu^2/2m$ we have

$$E_{\rm NR} = \frac{p^2}{2m} - \frac{\mu^2}{2m} \pm \frac{eB}{2m} \pm T. \tag{63}$$

From (42) it is seen that we should take $\mu = \frac{1}{2}{\rm Tr}K$; however, this gives an answer which is not in agreement with the nonrelativistic case. This does not come as any great suprise, for two reasons. First, there is no reason to expect that the limit $x^3 \to 0$ would commute with the nonrelativistic limit. It is impossible to work to consistent order when one has already set x^3 to zero. Second, in the nonrelativistic case we have neglected spin. Spin is easily added in the nonrelativistic formulation so that the covariant derivative is modified by a spin connection and the wave functions become two component wave functions. However, no terms arise in our effective mass or potential as a result of the spin connection. Thus we have that $\mu = \frac{1}{2}{\rm Tr}K$ and

$$\frac{\mu^2}{2m} = \frac{1}{2m}\frac{1}{4}{\rm Tr}K^2 \tag{64}$$

in partial agreement with (21). The missing terms in K cannot be corrected for unless one takes the nonrelativistic limit of the Dirac equation before separation, which is the correct procedure.

In summary, we have considered the reduction from 3+1 dimensions to 2+1 dimensions of a system of constrained fermions. Relativistic effects are considered. We find limits on the possiblie strength of a confining vector potential beyond which effective confinement breaks down. We show that surfaces which do not contain holes or conical singularities do not generate parity breaking theories when reducing from 3+1 dimensions to 2+1 dimensions.

APPENDIX: CONVENTIONS

Our conventions for the curvature follow those of Misner Thorne and Wheeler. The metric signature is $-+++$ and η_{ab} is the metric in flat Minkowski space-time. Relative to a vielbein field we define the Dirac matrices in 3+1 dimensions to satisfy

$$\{\gamma^a, \gamma^b\} = 2\eta^{ab} \to -(\gamma^0)^2 = (\gamma^i)^2 = 1. \tag{A1}$$

γ_5 is taken as

$$\gamma_5 = -\gamma^0\gamma^1\gamma^2\gamma^3 \tag{A2}$$

and

$$\{\gamma^a\,,\,\gamma_5\} = 0. \tag{A3}$$

The Pauli matrices are chosen so that

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In 2+1 dimensions the fundamental **2** representation of the γ matrices is satisfied by

$$\gamma^0 = -i\sigma_3, \quad \gamma^i = -\sigma_i \tag{A4}$$

for $i = 1, 2$. These satisfy

$$\gamma^a\gamma^b = \eta^{ab} + \epsilon^{abc}\gamma^c, \tag{A5}$$

$${\rm Tr}(\gamma^a\gamma^b\gamma^c) = -2\epsilon^{abc}, \tag{A6}$$

$$\gamma_5 = \gamma^0\gamma^1\gamma^2, \tag{A7}$$

$$[\gamma_5, \gamma_a] = 0. \tag{A8}$$

In the **4** representation in 2+1 dimensions the results are as for **2** except that

$${\rm Tr}(\gamma^a\gamma^b\gamma^c)_{\bf 4} = 0. \tag{A9}$$

[1] F. Wilczek, *Fractional Statistics and Anyon Superconductivity* (World Scientific, Singapore, 1990).
[2] E. Fradkin, *Field Theories of Condensed Matter Systems* (Addison-Wesley, New York, 1991).
[3] H. Jensen and H. Koppe, Ann. Phys. (N.Y.) **63**, 586, 1971.
[4] R.C. da Costa, Phys. Rev. A **23**, 1982 (1981).
[5] R.C. da Costa, Phys. Rev. A **25**, 2893 (1982).
[6] S. Matsutani, J. Phys. Soc. Jap. **61**, 55 1992).
[7] S. Matsutani, Phys. Rev. A **47**, 686 (1993).
[8] T. Ando, A. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
[9] T.H. Hansson, M. Roček, and I. Zahed, Phys. Lett. B **214**, 475 (1988).
[10] J.B. Marston and I. Affleck, Phys. Rev. B **39**, 11538 (1989).
[11] S.A. Fulling, *Aspects of Quantum Field Theory in Curved Spacetime* (Cambridge University Press, Cambridge, England, 1989).
[12] S.A. Fulling, Phys. Rev. D **14**, 1939 (1976).
[13] S. Matsutani and H. Tsuru, Phys. Rev. A **46**, 1144 (1992).
[14] A. Redlich, Phys. Rev. D **29**, 2366 (1984).
[15] J.M. Leinaas and J. Myrheim, Nuovo Cimento B **37**, 1 (1977).
[16] F. Wilczek, Phys. Rev. Lett. **48**, 1144 (1982).
[17] F. Wilczek and A. Zee, Phys. Rev. Lett. **51**, 2250 (1983).
[18] C. Manogue, Ann. Phys. (N.Y.) **181**, 261 (1988).
[19] A. Hansen and F. Ravndal, Phys. Scr. **23**, 1036 (1981).
[20] T. Damour, in *Proceedings from the First Marcel Grossmann Meeting on General Relativity*, edited by R. Ruffini (North-Holland, Amsterdam, 1977).
[21] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982).
[22] P. Exner and P. Seba, J. Math. Phys. **30**, 2574 (1989).
[23] E. Flekkøy and J.M. Leinaas, Int. J. Mod. Phys. A **6**, 5327 (1992).
[24] L. Ford, Phys. Rev. D **21**, 933 (1980).
[25] M. Burgess and D.J. Toms, Phys. Lett. B **252**, 596 (1990).