PHYSICAL REVIEW D VOLUME 26, NUMBER 2 15 JULY 1982

Eliminating lattice fermion doubling

Richard Stacey
*Department of Physics and Astronomy, University of College London, Gower Street, London WC1, England**
(Received 18 May 1981)

The problem of fermion doubling on a lattice is analyzed. Two new solutions are found in one space dimension. The first fails in the usual manner to generalize to three dimensions. The second solution generalizes to eliminate doubling in any number of dimensions.

I. INTRODUCTION

It has been known for some time that problems arise when we try to put massless fermions on a lattice. The usual formulation of the Weyl equation on a Kogut-Susskind lattice in $D+1$ dimensions describes not one but 2^D fermions.

Various procedures have been proposed for getting rid of the extra states.[1,2,3] None succeeds in eliminating them completely for $D=3$, except Ref. 3, which has problems with nonlocality. Now, in a quark model the doubling is not a disaster (e.g., Ref. 3), but it is in a theory of neutrinos. Recently, Nielsen and Ninomiya have argued[4] that doubling prevents our putting both weak interactions and chirally invariant quantum chromodynamics (QCD) on a lattice.

In this paper we argue that we can get around the conclusions of Ref. 4 and eliminate doubling altogether. In Sec. II we analyze the problem in one space dimension. Our analysis leads first to a partial solution (which does not generalize to three dimensions) and then to a full solution (Secs. III and IV).

II. DOUBLING IN ONE DIMENSION

A. The problem

Consider a massless fermion in $1+1$ dimensions. It obeys the Weyl equation

$$i\frac{\partial U}{\partial t}=i\alpha\frac{\partial U}{\partial x}\,, \tag{1}$$

where

$$U=\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$$

is a two-component spinor and

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

Let us put space on a lattice ($x=na$). It is usual to take

$$\frac{\partial U}{\partial x}\to\frac{1}{2a}[U(x+a)-U(x-a)]\,. \tag{2}$$

Substituting Eq. (2) in Eq. (1) Fourier transforming, we get

$$i\frac{\partial\tilde U(p)}{\partial t}=\left(\frac{\sin(pa)}{a}\alpha\right)\tilde U(p)\,. \tag{3}$$

The energy spectrum is given by the eigenvalues of the Hamiltonian matrix

$$H=\frac{\sin(pa)}{a}\alpha\,. \tag{4}$$

That is,

$$E^2=\frac{\sin^2(pa)}{a^2}\,. \tag{5}$$

Each zero eigenvalue corresponds to a massless fermion (e.g., Ref. 4), and there are *two* of these in the Brillouin zone p: $-\pi/a\to+\pi/a$, at $p=0$ *and* $p=\pi/a$. One fermion has become two. How?

B. Origin of doubling

The right-hand side of Eq. (3) is zero at $p=\pi/a$ because

$$\sum_n e^{i(\pi/a)na}[U(n+1)-U(n-1)]=0\,,$$

or

$$e^{-i(\pi/a)a}\sum_n e^{i(\pi/a)(n+1)a}U(n+1)=e^{+i(\pi/a)a}\sum_n e^{i(\pi/a)(n-1)a}U(n-1)\,,$$

or

$$e^{i(\pi/a)2a}=1\,. \tag{6}$$

The last is the crucial relation. *The problem arises because the lattice spacing (1a) is a factor of 2 less than the space between the points used to estimate ($\partial U/\partial x$) (2a).*

Another way to see this is to consider time-independent solutions. In the continuum Eq. (1) implies that these can only be given by a constant spinor. By contrast, if we use Eq. (2) on the lattice we find solutions

$$U(2n)=U(0)\forall n\,,\quad U(2m+1)=U(1)\forall m\,. \tag{7}$$

There is nothing in the theory to connect $U(0)$ and $U(1)$. The $2a$ spacing for $\partial U/\partial x$ distinguishes in this way between odd and even lattice sites, and gives rise to extra solutions. These last need not be continuous in the continuum limit.

In the general case we expand about the extra solution, taking

$$p\equiv\frac{\pi}{a}+p'\,. \tag{8}$$

Now

$$\begin{aligned} &e^{ipna}U(n)=e^{ip'na}[(-)^n U(n)]\,,\\ &\sin(pa)=-\sin(p'a)\,. \end{aligned} \tag{9}$$

If we treat p' as the momentum, therefore, the new solution is given by

$$U'(x)=(-)^n U(x)\quad (x=na) \tag{10}$$

and satisfies

$$i\frac{\partial\tilde{U}'(p')}{\partial t}=-\left(\frac{\sin(p'a)}{a}\alpha\right)\tilde{U}'(p')\,. \tag{11}$$

Owing to the lack of restriction on $U(x+a)-U(x)$, the space of solutions is too large. Every solution (U) is related to another (U') by a Z_2 transformation which leaves the theory invariant.

III. A PARTIAL SOLUTION

A. A new problem

We conclude from Sec. II that we need to make the lattice spacing the same as the distance over which $\partial U/\partial x$ is measured. Note that Susskind[3] effectively makes both $2a$ by separating ψ_1 and ψ_2 so that one lives on odd and the other on even lattice sites.

We will try to make the $\partial U/\partial x$ spacing just $1a$. The obvious method, replacing Eq. (2) by

$$\frac{\partial U}{\partial x}\to\frac{1}{a}\left[U(x+a)-U(x)\right], \tag{12}$$

fails immediately, since it leads to a non-Hermitian Hamiltonian with energies

$$E^2=e^{-ipa}\,\frac{4}{a^2}\sin^2\!\left(\frac{pa}{2}\right). \tag{13}$$

Clearly the same problem would arise if we used

$$\frac{\partial U}{\partial x}\to\frac{1}{a}\left[U(x)-U(x-a)\right]. \tag{14}$$

B. Solution A[5]

Instead of using Eq. (12) or (14) we can imitate Susskind[3] and distinguish between ψ_1 and ψ_2. We take

$$\begin{aligned} &\frac{\partial\psi_1}{\partial x}\to\frac{1}{a}[\psi_1(x+a)-\psi_1(x)]\,,\\ &\frac{\partial\psi_2}{\partial x}\to\frac{1}{a}[\psi_2(x)-\psi_2(x-a)]\,. \end{aligned} \tag{15}$$

Equation (15) leads to a Hermitian Hamiltonian

$$H=\begin{pmatrix} 0 & e^{+ipa/2}\dfrac{\sin(pa/2)}{a/2}\\ e^{-ipa/2}\dfrac{\sin(pa/2)}{a/2} & 0 \end{pmatrix} \tag{16}$$

with energy eigenvalues

$$E^2=\frac{4}{a^2}\sin^2\!\left(\frac{pa}{2}\right). \tag{17}$$

The energy spectrum has only one zero in the Brillouin zone, so there is only one massless fermion.

An alternative formulation of solution A is obtained if we define

$$\tilde{\psi}_2'(p)\equiv e^{ipa/2}\tilde{\psi}_2(p) \tag{18}$$

and

$$\tilde{U}'(p)\equiv\begin{pmatrix}\tilde{\psi}_1\\ \tilde{\psi}_2'\end{pmatrix}. \tag{19}$$

Now $\tilde{U}'(p)$ satisfies

$$i\frac{\partial\tilde{U}'(p)}{\partial t}=\frac{2}{a}\sin\!\left(\frac{pa}{2}\right)\alpha\tilde{U}'(p)\,. \tag{20}$$

Fourier transforming Eq. (18), we find that the transition

$$U\to U'$$

shifts ψ_2 from $x=na$ to $x=(n+\frac{1}{2})a$. ψ_1 and ψ_2 are now defined at alternate sites of a new lattice with spacing $\frac{1}{2}a$; and they satisfy Eq. (20). This is precisely as in the scheme of Susskind and Casher.[3] The only difference is that in the end we reinterpret everything on the $x=na$ lattice by

$$U'\to U\,.$$

C. Three space dimensions

In 3+1 dimensions the Weyl equation is

$$i\frac{\partial U(\vec{x})}{\partial t}=\frac{1}{i}\vec{\sigma}\cdot\vec{\partial}\,U(\vec{x})\,. \tag{21}$$

Putting space on a cubic lattice, the analog of Eq.

(2) gives an energy spectrum

$$E^2=\frac{1}{a^2}[\sin^2(p_1a)+\sin^2(p_2a)+\sin^2(p_3a)]\,. \qquad (22)$$

Equation (22) has $8=2^3$ zeros in the Brillouin zone, implying *eight* massless fermions.

The obvious generalization of solution A

$$\begin{aligned}\frac{\partial\psi_1}{\partial x_i}&\to\frac{1}{a}[\psi_1(\vec{x}+a\hat{x}_i)-\psi_1(\vec{x})]\,,\\ \frac{\partial\psi_2}{\partial x_i}&\to\frac{1}{a}[\psi_2(\vec{x})-\psi_2(\vec{x}-a\hat{x}_i)]\end{aligned} \qquad (23)$$

fails immediately. It gives a non-Hermitian Hamiltonian, as a result of the diagonal (σ_3) terms. Solutions in the manner of Susskind[3] also fail.

A possible outlet is suggested by another rephrasing of solution A in one dimension. That solution replaces the original Hamiltonian matrix $\{[\sin(pa)/a]\sigma_1\}$ by Eq. (16), which can be written

$$H=\frac{\sin(pa)}{a}\sigma_1-\left(\frac{1-\cos(pa)}{a}\right)\sigma_2\,. \qquad (24)$$

This is analogous to Wilson's use of

$$H=\frac{\sin(pa)}{a}\sigma_1+\frac{K}{a}[1-\cos(pa)]\sigma_3 \qquad (25)$$

in Ref. 1. Generalizing Eqs. (24) and (25) to three dimensions, we could take

$$H=\sum_i\left(\frac{\sin(p_ia)}{a}\sigma_i+\sum_{j=i}^{3}\frac{(1-\cos p_ja)}{a}\bar{K}_i^j\sigma_i\right) \qquad (26)$$

with the $\bar{K}^j$ real dimensionless three-vectors. Doing this we can easily reduce the number of fermions to two. Reduction to one is impossible. In three dimensions we have no extra degrees of freedom available to fill the role of σ_2 and σ_3 in Eqs. (24) and (25).

Solution A cannot be generalized to three dimensions without at least two fermions occurring.

IV. A FULL SOLUTION

A. One dimension

The problem with taking

$$\frac{\partial U}{\partial x}\to\frac{1}{a}[U(x+a)-U(x)] \qquad (12)$$

is that Eq. (12) naturally defines $\partial U/\partial x$ at $x+\frac{1}{2}a$. Thus we have U defined at $x=na$ ($\forall n$) and $\partial U/\partial x$ defined at $x=(m+\frac{1}{2})a$ ($\forall m$). In the Weyl equation both U and $\partial U/\partial x$ must be defined at the *same* point. Previously (in Sec. II) this was done by defining (at $x=na$)

$$\frac{\partial U}{\partial x}(n)\equiv\frac{1}{2}\left(\frac{\partial U}{\partial x}(n+\tfrac{1}{2})+\frac{\partial U}{\partial x}(n-\tfrac{1}{2})\right). \qquad (27)$$

However, we can equally well work at $x=(n+\frac{1}{2})a$, by averaging U itself:

$$U(n+\tfrac{1}{2})\equiv\tfrac{1}{2}[U(n)+U(n+1)]\,. \qquad (28)$$

Such a scheme gives our solution B.

First, we can show that the theory eliminates doubling, without reference to points $x=(n+\frac{1}{2})a$. We simply insist that the Weyl equation is satisfied using Eqs. (12) and (28). That is,

$$i\frac{\partial}{\partial t}\left(\frac{U(x+a)+U(x)}{2}\right)=i\alpha\left(\frac{U(x+a)-U(x)}{a}\right). \qquad (29)$$

Equation (29) Fourier transforms to

$$i\frac{\partial\tilde{U}}{\partial t}(p)=\frac{2}{a}\tan\left(\frac{pa}{2}\right)\alpha\tilde{U}(p)\,. \qquad (30)$$

The Hamiltonian matrix

$$\left[\frac{2}{a}\tan\left(\frac{pa}{2}\right)\alpha\right]$$

is Hermitian, with energy eigenvalues

$$E^2=\frac{4}{a^2}\tan^2\frac{pa}{2}\,. \qquad (31)$$

Equation (31) has the correct continuum limit, and only one zero in the Brillouin zone. The theory describes *just one massless fermion.*

Next we use the $\{U(n)\}$ to define a new wave function $V(x)$ and $\partial V/\partial x$ at $x=(n+\frac{1}{2})a$. We take

$$\begin{aligned}V(x)&\equiv\tfrac{1}{2}[U(x+\tfrac{1}{2}a)+U(x-\tfrac{1}{2}a)]\,,\\ \partial V/\partial x&\Rightarrow\frac{1}{a}[U(x+\tfrac{1}{2}a)-U(x-\tfrac{1}{2}a)]\end{aligned} \qquad (32)$$

[cf. Eqs. (12) and (28)]. Now V and $\partial V/\partial x$ must satisfy the Weyl equation, and we get Eqs. (30) and (31) again with

$$\tilde{V}=\cos\frac{pa}{2}\,\tilde{U}\,. \qquad (33)$$

Note that in the continuum limit V must be continuous and equal to U. The theory in terms of the V is given by a configuration-space Hamiltonian

$$\begin{aligned}H&=\sum_{(n+1/2)a}V^\dagger(-i\alpha)\frac{\partial V}{\partial x}\\ &=\sum_{na}\left(\frac{U(x)+U(x+a)}{2}\right)^\dagger(-i\alpha)\left(\frac{U(x+a)-U(x)}{a}\right)\,.\end{aligned} \qquad (34)$$

Equation (34) gives the same Hamiltonian as in Sec. II. The theories described are not the same since the fundamental fields are different (U and V).

B. Three dimensions

The above solution generalizes to three space dimensions without difficulty. Again, the theory is best pictured as using the $U(\vec{x})$ to define a new wave function $V(\vec{x})$ and its derivatives on a second cubic lattice with sites at

$$\vec{x}:\ (\vec{x})_i = (n_i + \tfrac{1}{2})a\ \forall i\,. \tag{35}$$

We define

$$V(\vec{x}) \equiv \tfrac{1}{8} \sum_{\substack{\pm \\ \text{all combinations}}} U\left(x_1 \pm \frac{a}{2}, x_2 \pm \frac{a}{2}, x_3 \pm \frac{a}{2}\right) \tag{36}$$

and

$$\frac{\partial V}{\partial x_1} = \frac{1}{4a}\sum_{\pm}\left[U\left(x_1+\frac{a}{2}, x_2 \pm \frac{a}{2}, x_3 \pm \frac{a}{2}\right) - U\left(x - \frac{a}{2}, x_2 \pm \frac{a}{2}, x_3 \pm \frac{a}{2}\right)\right] \tag{37}$$

with $\partial V/\partial x_2$ and $\partial V/\partial x_3$ defined similarly.

Substituting Eqs. (36) and (37) in the Weyl equation and Fourier transforming, we get

$$i\frac{\partial \tilde{V}(p)}{\partial t} = -\left[\sum_i \frac{2}{a}\tan\left(\frac{p_i a}{2}\right)\sigma_i\right]\tilde{V}(p)\,. \tag{38}$$

The Hamiltonian matrix is Hermitian, with an energy spectrum given by

$$E^2 = \frac{4}{a^2}\left(\tan^2\frac{p_1 a}{2} + \tan^2\frac{p_2 a}{2} + \tan^2\frac{p_3 a}{2}\right)\,. \tag{39}$$

There is only one zero in the Brillouin zone, so the theory describes just one massless fermion. The doubling problem has disappeared.

Note that (as in one dimension) we can express the theory in terms of the $U(x)$ alone, without reference to $V(x)$ or the new lattice. Then

$$\tilde{U} = \prod_{i=1}^{3}\left(\sec\frac{p_i a}{2}\right)\tilde{V} \tag{40}$$

satisfies Eq. (38), and gives rise to Eq. (39).

The configuration-space Hamiltonian [which only makes sense in terms of the $V(x)$] is

$$H = \sum_{x,i} V^\dagger i\sigma_i \frac{\partial V}{\partial x_i} = \sum U(x) i\sigma_i \frac{\partial U'}{\partial x_i}\ , \tag{41}$$

where

$$\begin{aligned}\frac{\partial U'}{\partial x_i} \equiv \frac{1}{32a}\Big\{&4[U(x_1+a,x_2,x_3) - U(x_1-a,x_2,x_3)]\\
&+2[U(x_1+a,x_2+a,x_3) - U(x_1-a,x_2+a,x_3)]\\
&+2[U(x_1+a,x_2-a,x_2) - U(x_1-a,x_2-a,x_3)]\\
&+2[U(x_1+a,x_2,x_3+a) - U(x_1-a,x_2,x_3+a)]\\
&+2[U(x_1+a,x_2,x_3-a) - U(x_1-a,x_2,x_3-a)]\\
&+\sum_{\pm}[U(x_1+a,x_2\pm a,x_3\pm a)\\
&\quad - U(x_1-a,x_2\pm a,x_3\pm a)]\Big\}.\end{aligned} \tag{42}$$

Note that Eq. (42) gives an acceptable, if unconventional, approximation to $\partial U/\partial x$ in the standard theory.

C. Some comments on the solution

(1) Solution B does not belong to the class of theories considered in Ref. 4, because neither the Hamiltonian nor the individual space derivatives $(\partial V/\partial x)$ can be expressed as a combination of the wave functions $V(x)$. One way to see this is to note that if we were to define a theory of W spinors at $y_n = (n+\frac{1}{2})a$, such that

$$U(n) = \tfrac{1}{2}\,[W(n+\tfrac{1}{2}) + W(n-\tfrac{1}{2})]\,,$$

we could choose (say) $W(\frac{1}{2})$ to have any value at all. The $\{W(x)\}$ have one more degree of freedom than the $\{U(x)\}$. For this reason the conclusions of Ref. 4 do not apply.

(2) The zero of energy at $p=\pi/a$ has been replaced by a pole (1 dimension). This means that the momentum is restricted to a finite range (e.g., $-\pi/a \to \pi/a$) but the energy can take all values in that range.

(3) The fact that $U(x)$ is continuous in the continuum limit distinguishes the new class of solutions from that considered in Sec. II. Here the only time-independent solutions are

$$U(x) = U(0)\forall x\,.$$

(4) We *could* define $V(x)$ on the old lattice sites by

$$V'(x) = V\left(x_1+\frac{a}{2}, x_2+\frac{a}{2}, x_3+\frac{a}{2}\right).$$

It would give essentially the same theory, in a less natural way. More natural is to consider solution B as implementing the Weyl equation on the d cubes of the lattice, rather than at the points.

(5) Solution B generalizes to eliminate doubling in any number of dimensions.

V. CONCLUSIONS

In Sec. II we analyzed the fermion lattice doubling problem in $1+1$ dimensions. Our analysis led first to solution A (Sec. III) which failed to generalize to three dimensions for the usual reasons (cf. Refs. 1 and 3). Note that like the solutions of Refs. 1 and 3, our solution A is not chirally invariant, so the results of Ref. (4) are not relevant.

Solution B (Sec. IV) was rather different. We chose to redefine the wave functions, as well as their space derivatives, for use in the Weyl equation. The new theory is best pictured as using the $\{U(x)\}$ to define (in a natural way) new wave functions and their space derivatives on a new lattice with sites at the center of the cubes of the first.

Solution B eliminates doubling in any number of dimensions. Using it we should be able to put both weak interactions and chirally invariant QCD models on the lattice, in spite of the negative results of Ref. 4.

ACKNOWLEDGMENT

I would like to thank the United Kingdom Science Research Council for financial support.

APPENDIX: IS THE THEORY LOCAL?

The theory as defined by Eq. (32) (in one dimension) appears local. However, if we write

$$\frac{\partial V}{\partial x}(x)=\frac{\lambda}{a}\left(\sum_{n=1}^{\infty}(-)^{n+1}[V(x+na)-V(x-na)]\right), \quad \text{(A1)}$$

it appears nonlocal. The infinite series in (A1) results from the pole in $\tan(pa/2)$ at $p=\pi/a$. Such a nonlocal resolution of doubling would be analogous to that of Ref. 2.

Here I will argue that the nonlocality apparent in Eq. (A1) is deceptive, because it ignores the underlying $U(x)$ structure. Basically, we cannot change one $V(x+na)$ in Eq. (A1)—$n\geq 2$—without changing others in such a way that the nonlocal variations cancel. More precisely, for $n\geq 2$ any $V(x+na)$ in Eq. (A1) occurs in a combination of the form

$$\cdots - V(x+(n-1)a)+V(x+na)-V(x+(n+1)a)+\cdots .$$

Recalling that

$$V(x+na)=\tfrac{1}{2}\left[U(x+(n-\tfrac{1}{2})a)+U(x+(n+\tfrac{1}{2})a)\right] \quad \text{(A2)}$$

we see that the $U(x+(n\pm\frac{1}{2})a)$ term in $V(x+na)$ cancels with the corresponding term in $-V(x+(n\pm 1)a)$, i.e., both terms cancel. This argument suggests that the theory really is local.

---

*Now at DAMTP, P. O. Box 147, Liverpool L69, England.

[1]K. G. Wilson, in *New Phenomena in Subnuclear Physics*, proceedings of the 14th Course of the International School of Subnuclear Physics, Erice, 1975, edited by A. Zichichi (Plenum, New York, 1977).

[2]S. D. Drell, M. Weinstein, and S. Yankielowicz, Phys. Rev. D 14, 489 (1976).

[3]L. Susskind, Phys. Rev. D 16, 3031 (1977).

[4]H. B. Nielsen and M. Ninomiya, Nucl. Phys. B185, 20 (1981).

[5]It has been brought to my attention that our solution A has been treated previously in Y. Nakawaki, Prog. Theor. Phys. 61, 1855 (1979).