RENORMALIZATION OF THE REGULARIZED RELATIVISTIC ELECTRON-POSITRON FIELD

ELLIOTT H. LIEB AND HEINZ SIEDENTOP

Abstract. We consider the relativistic electron-positron field interacting with itself via the Coulomb potential defined with the physically motivated, positive, density-density quartic interaction. The more usual normal-ordered Hamiltonian differs from the bare Hamiltonian by a quadratic term and, by choosing the normal ordering in a suitable, self-consistent manner, the quadratic term can be seen to be equivalent to a renormalization of the Dirac operator. Formally, this amounts to a Bogolubov-Valatin transformation, but in reality it is non-perturbative, for it leads to an inequivalent, fine-structure dependent representation of the canonical anticommutation relations. This non-perturbative redefinition of the electron/positron states can be interpreted as a mass, wave-function and charge renormalization, among other possibilities, but the main point is that a non-perturbative definition of normal ordering might be a useful starting point for developing a consistent quantum electrodynamics.

1. Introduction

In relativistic quantum electrodynamics (QED) the quantized electron-positron field $\Psi(x)$, which is an operator-valued spinor, is written formally as

$$\Psi(x) := a(x) + b^*(x), \tag{1}$$

where $a(x)$ annihilates an electron at x and $b^*(x)$ creates a positron at x. (We use the notation that x denotes a space-spin point, namely $x = (\mathbf{x}, \sigma) \in \Gamma = \mathbb{R}^3 \times \{1,2,3,4\}$ and $\int d^3x$ denotes integration over $\mathbb{R}^3$ and a summation over the spin index.) More precisely, we take the Hilbert space, $\mathfrak{H}$ of $L^2(\mathbb{R}^3)$ spinors, i.e., $\mathfrak{H} = L^2(\mathbb{R}^3) \otimes \mathbb{C}^4$. To specify the one-electron space we choose a subspace, $\mathfrak{H}_+$ of $\mathfrak{H}$ and denote the orthogonal projector onto this subspace by P_+. The one-positron space is the (anti-unitary) charge conjugate, $C\mathfrak{H}_-$, of the orthogonal complement, $\mathfrak{H}_-$ with projector P_-. In position space, $(C(\psi))(\mathbf{x}) := i\beta\alpha_2\overline{\psi(\mathbf{x})}$, whereas in momentum space $(C(\hat\psi))(\mathbf{p}) := i\beta\alpha_2\overline{\hat\psi(-\mathbf{p})}$. For later purposes we note explicitly that the Hilbert space $\mathfrak{H}$ can be written as the orthogonal sum

$$\mathfrak{H} = \mathfrak{H}_+ \oplus \mathfrak{H}_-. \tag{2}$$

If f_ν is an orthonormal basis for $\mathfrak{H}_+$ with $\nu \geq 0$ and an orthonormal basis for $\mathfrak{H}_-$ with $\nu < 0$ then $a(x) := \sum_{\nu=0}^{\infty} a(f_\nu) f_\nu(x)$, and $b(x) := \sum_{\nu=-\infty}^{-1} b(f_\nu)\overline{f_\nu(x)}$, where $a^*(f)$ creates an electron in the state $P_+ f$, etc. (For further details about the notation see Thaller [6], and also Helffer and Siedentop [4] and Bach et al [2, 1].

Date: March 2, 2000.

Key words and phrases. Dirac operator, ground state energy, quantum electrodynamics, regularization, renormalization.

For free electrons and positrons, $\mathfrak{H}_+$ and $\mathfrak{H}_-$ are, respectively, the positive and negative energy solutions of the free Dirac operator

$$D_0 = \boldsymbol{\alpha}\cdot\mathbf{p} + m_0\beta, \tag{3}$$

in which $\boldsymbol{\alpha}$ and β denote the four 4×4 Dirac matrices and $\mathbf{p} = -i\nabla$. The number m_0 is the *bare mass* of the electron/positron. Perturbation theory is defined in terms of this splitting of $\mathfrak{H}$ into $\mathfrak{H}_+$ and $\mathfrak{H}_-$. The electron/positron lines in Feynman diagrams are the resolvents of D_0 split in this way. As we shall see, this splitting may not be the best choice, ultimately, and another choice (which will require a fine-structure dependent, inequivalent representation of the canonical anticommutation relations) might be more useful.

The Hamiltonian for the free electron-positron field is

$$\mathbb{H}_0 = \int d^3x : \Psi(x)^* D_0 \Psi(x) :$$

with this choice of $\mathfrak{H}_+$. The symbol $: :$ denotes normal ordering, i.e., anti-commuting all a^*, b^* operators to the left (but ignoring the anti-commutators).

In this paper we investigate the effect of the Coulomb interaction among the particles, which is a quartic form in the $a^\#, b^\#$ operators. ($\#$ denotes either a star or no star.) The normal ordering will give rise to extra constants and certain quadratic terms. Our main point will be that by making an appropriate choice of $\mathfrak{H}_+$ that is different from the usual one mentioned above, we can absorb the additional quadratic terms into a mass, wave function, and charge renormalization. This choice of $\mathfrak{H}_+$ has to be made self consistently and, as we show, can be successfully carried out if some combination of the fine structure constant α (=1/137 in nature) and the ultraviolet cutoff Λ is small enough (but 'small' is actually 'huge' since the condition is, roughly, $\alpha \log\Lambda < 1$).

The physical significance of our construction is not immediate, but it does show that certain annoying terms, when treated *non-perturbatively*, can be incorporated into the renormalization program. It is not at all clear that what we have called wave function renormalization, for example, really corresponds to a correct interpretation of that renormalization, or even what the true meaning of wave function renormalization is. It could also be interpreted as a renormalization of Planck's constant. Likewise, the charge renormalization we find is not mandatory. But our main point is that the effect is mathematically real and is best dealt with by a change of the meaning of the one-particle electron/positron states. It should be taken into account in the non-perturbative QED that is yet to be born. Of course, there are other renormalization effects in QED that we do not consider. In particular, the magnetic field has not been included and so we do not have a Ward identity to help us. Our result here is only a small part of the bigger picture of a proper, non-perturbative QED.

The last section of this paper contains a brief discussion of some possible interpretations of our findings and the interested reader is urged to look at that section.

Our starting point is the unrenormalized ('bare') Hamiltonian for a quantized electron/positron field interacting with itself via Coulomb's law, namely

$$\begin{aligned}(4)\quad &\mathbb{H}^{\mathrm{bare}}\\ &:= \int d^3x : \Psi^*(x) D_0 \Psi(x) : + \frac{\alpha}{2}\int d^3x \int d^3y W(\mathbf{x},\mathbf{y}) : \Psi^*(x)\Psi(x) :: \Psi^*(y)\Psi(y) : \ .\end{aligned}$$

where W is a symmetric ($W(x,y) = W(y,x)$) interaction potential. The case of interest here is the Coulomb potential $W(x,y) = \delta_{\sigma,\tau}|\mathbf{x}-\mathbf{y}|^{-1}$ and regularized versions thereof. We will refer to the first term on the right side of (4) as the kinetic energy $\mathbb{T}$ and to the second term as the interaction energy $\mathbb{W}^{\text{bare}}$.

Our choice of the product of the two normal ordered factors in the interaction, namely $:\rho(x)::\rho(y):$, is taken from the book of Bjorken and Drell [3] (Eq. (15.28) without magnetic field). It is quite natural, we agree, to start with the closest possible analog to the classical energy of a field. Of course, this term is only partly normal ordered in the usual sense and therefore it does not have zero expectation value in the unperturbed vacuum or one-particle states. It is positive, however, as a Coulomb potential ought to be.

The fully normal ordered interaction (which is not positive) is defined to be

$$\mathbb{W}_\alpha^{\text{ren}} = \frac{\alpha}{2}\int d^3x \int d^3y W(\mathbf{x},\mathbf{y}) : \Psi^*(x)\Psi^*(y)\Psi(y)\Psi(x) : . \tag{5}$$

Bear in mind that this definition entails a determination of the splitting (2) of $\mathfrak{H}$.

We introduce a normal ordered ('dressed') Hamiltonian

$$\mathbb{H}_\alpha^{\text{ren}} := \int d^3x : \Psi^*(x) D_{Z,m}\Psi(x) : + \mathbb{W}^{\text{ren}\,\alpha}. \tag{6}$$

(Again, this depends on the splitting of $\mathfrak{H}$.) The operator $D_{Z,m}$ differs from D_0 in two respects:

$$D_{Z,m} = Z^{-1}(\boldsymbol{\alpha}\cdot\mathbf{p} + m\beta), \tag{7}$$

where $Z < 1$ and $m > m_0$. We call m_0 the bare mass and m the 'physical' mass — up to further renormalizations, which we do not address in this paper. The factor Z is called the 'wave function renormalization' because the 'wave function renormalization' in the standard theory is defined by the condition that the electron-positron propagator equals Z times a non-interacting propagator, which is the inverse of a Dirac operator. To the extent that $\mathbb{W}_\alpha^{\text{ren}}$ can be neglected, the propagator with just the first term on the right side of (6) would, indeed, be Z times a Dirac operator with mass m. One could also interpret Z in other ways, e.g., as a renormalization of the speed of light, but such a choice would be considerably more radical. Some other interpretations are discussed in the last section.

Our goal is to show that by a suitable choice of Z, m, and the electron space $\mathfrak{H}_+$, we have – up to physically unimportant infinite additive constants – asymptotic equality between $\mathbb{H}^{\text{bare}}$ and $\mathbb{H}_\alpha^{\text{ren}}$ for momenta that are much smaller than the cutoff Λ.

$$\mathbb{H}^{\text{bare}} = \mathbb{H}_\alpha^{\text{ren}} \qquad \text{(for small momenta)} \tag{8}$$

The additional normal ordering in (6) will introduce some quadratic terms and it is these terms to which we turn our attention and which we would like to identify as renormalization terms. Whether our picture is really significant physically, or whether it is just a mathematical convenience remains to be seen. What it does do is indicate a need for starting with a non-conventional, non-perturbative choice of the free particle states, namely the choice of $\mathfrak{H}_+$.

2. Calculation of the New Quadratic Terms

The canonical anti-commutation relations are

$$\begin{aligned}\{a(f),a(g)\}&=\{a^*(f),a^*(g)\}=\{a(f),b(g)\}=\{a^*(f),b^*(g)\}\\&=\{a^*(f),b(g)\}=\{a(f),b^*(g)\}=0,\end{aligned}\tag{9}$$

and

$$\{a(f),a^*(g)\}=(f,P_+g),\ \{b^*(f),b(g)\}=(f,P_-g).\tag{10}$$

Formally, this is equivalent to

$$\begin{aligned}\{a(x),a(y)\}&=\{a^*(x),a^*(y)\}=\{a(x),b(y)\}=\{a^*(x),b^*(y)\}\\&=\{a^*(x),b(y)\}=\{a(x),b^*(y)\}=0,\end{aligned}\tag{11}$$

and

$$\{a(x),a^*(y)\}=P_+(x,y),\ \{b^*(x),b(y)\}=P_-(x,y),\tag{12}$$

where $P_\pm(x,y)$ is the integral kernel of the projector $P_\pm$.

In order to compare the bare Hamiltonian $\mathbb{H}^{\mathrm{bare}}$ with the renormalized one $\mathbb{H}^{\mathrm{ren}}_\alpha$ we must face a small problem; the momentum cutoff Λ, which is necessary in order to get finite results, will spoil the theory at momenta comparable to this cutoff. Thus, our identification of the difference as a wave function and mass renormalization will be exact only for momenta that are small compared to Λ. We will also require (although it is not clear if this is truly necessary or whether it is an artifact of our method) that $\alpha\ln(\Lambda/m_0)$ is not too large.

The potential energy $\mathbb{W}^{\mathrm{bare}}$ clearly has positive singularities in it if the unregularized Coulomb interaction $W(x,y)=|\mathbf{x}-\mathbf{y}|^{-1}\delta_{\sigma,\tau}$ is taken. We therefore cut off the field operators for high momenta above Λ. (Another choice would be to cut off the Fourier transform of $|\mathbf{x}-\mathbf{y}|^{-1}$ at Λ and the result would be essentially the same; our choice is motivated partly by computational convenience.) Our cutoff procedure is to require that

$$\Psi(x):=(2\pi)^{-3/2}\int_{\{\mathbf{p}\in\mathbb{R}^3\ |\ |\mathbf{p}|<\Lambda\}}\widehat{\Psi}(\mathbf{p},\sigma)e^{-i\mathbf{p}\cdot\mathbf{x}}d\mathbf{p}\tag{13}$$

holds, with $\Lambda>0$. We could also take a smooth cutoff without any significant change. This regularization keeps the positivity of $\mathbb{W}^{\mathrm{bare}}$. To make all terms finite we also need a volume cutoff for one of the difference terms. However, the volume singularity only occurs as an additive constant, which we drop. The renormalized Hamiltonian is free of infrared divergent energies and so does not need a volume cutoff.

Next, we calculate the difference

(14)
$$\begin{aligned}\mathbb{R}:=\mathbb{W}^{\mathrm{bare}}-\mathbb{W}^{\mathrm{ren}}_\alpha=\frac{\alpha}{2}\int d^3x\int d^3y\ W(\mathbf{x},\mathbf{y})[&a^*(x)a(y)P_+(x,y)+a^*(x)b^*(y)P_+(x,y)\\&+b(x)a(y)P_+(x,y)-b^*(y)b(x)P_+(x,y)\\&-a^*(y)a(x)P_-(x,y)+P_-(x,y)P_+(x,y)+a(x)b(y)P_-(x,y)\\&+b^*(x)a^*(y)P_-(x,y)+b(x)^*b(y)P_-(x,y)]\ .\end{aligned}$$

Thus, (4) can be rewritten as

$$
\begin{aligned}
(15)\quad \mathbb{H}^{\mathrm{bare}} = &\int d^3x : \Psi^*(x) D_0 \Psi(x) : \\
&+ \frac{\alpha}{2} \int d^3x \int d^3y \frac{P_+(x,y) - P_-(x,y)}{|\mathbf{x}-\mathbf{y}|} : \Psi^*(x)\Psi(y) : + \mathbb{W}_\alpha^{\mathrm{ren}} \\
&\qquad + \frac{\alpha}{2} \int d^3x \int d^3y \frac{P_+(x,y)P_-(x,y)}{|\mathbf{x}-\mathbf{y}|}.
\end{aligned}
$$

The first two terms on the right side are one-particle operators (i.e., quadratic in the field operators). Let us call their sum $\mathbb{A}$, which we write (formally, since A is a differential operator) as

$$\mathbb{A} = \int d^3x \int d^3y A(x,y) : \Psi^*(x)\Psi(y) : . \tag{16}$$

The last term is a cutoff dependent constant, which happens to be infinite. (One can show that this term is positive and, if we put in a momentum cutoff Λ and volume (infrared) cutoff V, this integral is proportional to $V\Lambda^4$.)

We can make $\mathbb{A}$ positive by choosing the normal ordering appropriately. That is, we choose P_+ to be the projector onto the positive spectral subspace of the operator A. Thus, we are led to the nonlinear equation for the unknown A

$$A = D_0 + \frac{\alpha}{2} R, \tag{17}$$

where the operator R has the integral kernel

$$R(x,y) := \frac{P_+(x,y) - P_-(x,y)}{|\mathbf{x}-\mathbf{y}|} = \frac{(\operatorname{sgn} A)(x,y)}{|\mathbf{x}-\mathbf{y}|}, \tag{18}$$

with $\operatorname{sgn} t$ being the sign of t.

For physical reasons and to simplify matters we restrict the search for a solution to (17) to translationally invariant operators, i.e., 4×4 matrix-valued Fourier multipliers. Moreover, we make the ansatz

$$A(\mathbf{p}) := \boldsymbol{\alpha} \cdot \omega_{\mathbf{p}} g_1(|\mathbf{p}|) + \beta g_0(|\mathbf{p}|) \tag{19}$$

with real functions g_1 and g_0, where $\omega_{\mathbf{p}}$ is the unit vector in the $\mathbf{p}$ direction. In other words, we try to make A look as much as possible like a Dirac operator. With this ansatz we have $\operatorname{sgn} A(\mathbf{p}) = A(\mathbf{p})/(g_1(|\mathbf{p}|)^2 + g_0(|\mathbf{p}|)^2)^{1/2}$. Thus, recalling (13), (17) can be fulfilled, if

$$g_0(|\mathbf{p}|) := m_0 + \frac{\alpha}{4\pi^2} \int_{|\mathbf{q}|<\Lambda} d\mathbf{q} \frac{1}{|\mathbf{p}-\mathbf{q}|^2} \frac{g_0(|\mathbf{q}|)}{(g_1(|\mathbf{q}|)^2 + g_0(|\mathbf{q}|)^2)^{1/2}} \tag{20}$$

$$g_1(|\mathbf{p}|) := |\mathbf{p}| + \frac{\alpha}{4\pi^2} \int_{|\mathbf{q}|<\Lambda} d\mathbf{q} \frac{\omega_{\mathbf{p}} \cdot \omega_{\mathbf{q}}}{|\mathbf{p}-\mathbf{q}|^2} \frac{g_1(|\mathbf{q}|)}{(g_1(|\mathbf{q}|)^2 + g_0(|\mathbf{q}|)^2)^{1/2}} \tag{21}$$

is solvable. Note that the bare mass m_0 appears in (20).

3. Determination of the Dressed Electron

We will find a solution of the system (20), (21) by a fixed point argument. To this end we first integrate out the angle on the right hand side. Setting $u := |\mathbf{p}|$

and $v := |\mathbf{q}|$ we get

$$g_0(u) = m_0 + \frac{\alpha}{2\pi}\int_0^\Lambda dv \frac{v}{u} Q_0\left(\frac{1}{2}\left(\frac{u}{v}+\frac{v}{u}\right)\right)\frac{g_0(v)}{(g_1(v)^2+g_0(v)^2)^{1/2}} \tag{22}$$

$$g_1(u) = u + \frac{\alpha}{2\pi}\int_0^\Lambda dv \frac{v}{u} Q_1\left(\frac{1}{2}\left(\frac{u}{v}+\frac{v}{u}\right)\right)\frac{g_1(v)}{(g_1(v)^2+g_0(v)^2)^{1/2}} \tag{23}$$

where, for $z > 1$,

$$Q_0(z) = \frac{1}{2}\log\frac{z+1}{z-1} \tag{24}$$

and

$$Q_1(z) = \frac{z}{2}\log\frac{z+1}{z-1} - 1. \tag{25}$$

The case of zero bare mass is particularly easy. We can choose $g_0 = 0$, in which case g_1 is obtained by integration. We get

$$\begin{aligned} g_1(u) =& u + \frac{\alpha}{2\pi}\int_0^\Lambda dv\frac{v}{u}Q_1\left(\frac{1}{2}\left(\frac{u}{v}+\frac{v}{u}\right)\right)\\ =& u + \frac{\alpha}{2\pi}u\int_0^{\Lambda/u} dv\, v\, Q_1\left(\left(1/v+v\right)/2\right)\\ =& u + \frac{\alpha}{2\pi}u\left[\frac{2}{3}\log|(\Lambda/u)^2-1| - \frac{1}{3}(\Lambda/u)^2\right.\\ &\left.+\frac{\Lambda}{6u}\log\left|\frac{(\Lambda/u)+1}{(\Lambda/u)-1}\right|\left(3+(\Lambda/u)^2\right)\right] \end{aligned} \tag{26}$$

The function g_1 behaves asymptotically for small u or large Λ as $g_1(u) = \frac{\alpha}{2\pi}\frac{4}{3}u\log(\Lambda/u)$. Unfortunately, because of the $\log u$, the operator A can never look like the renormalized Dirac operator in (7). This asymptotic expansion can be seen either from (26) or else by noting that the integrand is a continuous function, except for $v = 1$, that decays at infinity as $4/(3v)$. This follows from the large v expansion

$$vQ_1\left((1/v+v)/2\right) = \frac{4}{3}v^{-1} + \frac{8}{15}v^{-3} + O(v^{-5}). \tag{27}$$

For positive bare masses we solve the system (22) and (23) by a fixed point argument. To this end we define the following set of pairs of functions, for $\epsilon, \delta > 0$.

$$S_{\epsilon,\delta} := \{\mathbf{g} = (g_0, g_1)\,|\,\forall_{u\in[0,\Lambda]}\mathbf{g}(u) \in [m_0, (1+\delta)m_0]\times[u,(1+\epsilon)u]\}. \tag{28}$$

Note that with the metric generated by the sup norm $S_{\epsilon,\delta}$ is a complete metric space. Next define $T : S_{\epsilon,\delta} \to S_{\epsilon,\delta}$ by the right hand side of (23) and (22), i.e.,

$$T_0(\mathbf{g})(u) = m_0 + \frac{\alpha}{2\pi}\int_0^\Lambda dv \frac{v}{u} Q_0\left(\frac{1}{2}\left(\frac{u}{v}+\frac{v}{u}\right)\right)\frac{g_0(v)}{(g_1(v)^2+g_0(v)^2)^{1/2}} \tag{29}$$

$$T_1(\mathbf{g})(u) = u + \frac{\alpha}{2\pi}\int_0^\Lambda dv \frac{v}{u} Q_1\left(\frac{1}{2}\left(\frac{u}{v}+\frac{v}{u}\right)\right)\frac{g_1(v)}{(g_1(v)^2+g_0(v)^2)^{1/2}}. \tag{30}$$

Lemma 1. *If Y, defined by*

$$Y := \alpha\,\mathrm{arsinh}(\Lambda/m_0)/\pi, \tag{31}$$

satisfies $Y < 9/50$, if $\epsilon \geq 50Y/(9-50Y)$, and if $\delta \geq Y/(1-Y)$, then T maps $S_{\epsilon,\delta}$ into $S_{\epsilon,\delta}$.

Note that $\operatorname{arsinh}(x) = \log(x + \sqrt{x^2+1})$, i.e., Y grows logarithmically in Λ/m_0.

Proof. Obviously, $T_0(\mathbf{g})(u) \geq m_0$ and $T_1(\mathbf{g})(u) \geq u$.

To bound g_0 we return to (20) and note the bound (which holds in $S_{\epsilon,\delta}$)

$$
\begin{aligned}
g_0(|\mathbf{p}|) &\leq m_0 + \frac{\alpha}{4\pi^2}\int_{|\mathbf{q}|<\Lambda} d\mathbf{q}\frac{1}{|\mathbf{p}-\mathbf{q}|^2}\frac{(1+\delta)m_0}{(|\mathbf{q}|^2+m_0^2)^{1/2}} \\
&\leq m_0 + \frac{\alpha}{4\pi^2}\int_{|\mathbf{q}|<\Lambda} d\mathbf{q}\frac{1}{\mathbf{q}^2}\frac{(1+\delta)m_0}{(|\mathbf{q}|^2+m_0^2)^{1/2}} \\
&= m_0[1+\frac{\alpha}{\pi}(1+\delta)\operatorname{arsinh}(\Lambda/m_0)].
\end{aligned} \tag{32}
$$

The second inequality holds because the convolution of two symmetric decreasing functions is symmetric decreasing.

Next we turn to g_1, which is a bit more complicated. We split the integration region in (21) into $A = \{\mathbf{q} \,|\, |\mathbf{q}| < 2|\mathbf{p}| \text{ and } |\mathbf{q}| < \Lambda\}$ and $B = \{\mathbf{q} \,|\, |\mathbf{q}| \geq 2|\mathbf{p}| \text{ and } |\mathbf{q}| < \Lambda\}$. In region A we use $\omega_{\mathbf{p}} \cdot \omega_{\mathbf{q}} \leq 1$ and hence the contribution to g_1 from this region is bounded above by

$$
\frac{\alpha}{4\pi^2}\int_{|\mathbf{q}|<\Lambda} d\mathbf{q}\frac{1}{|\mathbf{p}-\mathbf{q}|^2}\frac{(1+\epsilon)2|\mathbf{p}|}{(|\mathbf{q}|^2+m_0^2)^{1/2}} \leq 2|\mathbf{p}|\frac{\alpha}{\pi}(1+\epsilon)\operatorname{arsinh}(\Lambda/m_0)], \tag{33}
$$

for the same reason as in (32).

In region B we use that $(|\mathbf{p}|^2 - |\mathbf{q}|^2)^2 \geq 9|\mathbf{q}|^4/16$. We also note, for the integration over B, that we can take the mean of the integrand for $\mathbf{q}$ and $-\mathbf{q}$. In other words, we can bound this term as follows:

$$
\begin{aligned}
&\frac{\alpha}{4\pi^2}\left|\int_B d\mathbf{q}\frac{\omega_{\mathbf{p}}\cdot\omega_{\mathbf{q}}}{|\mathbf{p}-\mathbf{q}|^2}\frac{g_1(|\mathbf{q}|)}{(g_1(|\mathbf{q}|)^2+g_0(|\mathbf{q}|)^2)^{1/2}}\right| \\
\leq&\frac{\alpha}{8\pi^2}\left|\int_B d\mathbf{q}\left(\frac{\omega_{\mathbf{p}}\cdot\omega_{\mathbf{q}}}{|\mathbf{p}-\mathbf{q}|^2} - \frac{\omega_{\mathbf{p}}\cdot\omega_{\mathbf{q}}}{|\mathbf{p}+\mathbf{q}|^2}\right)\frac{g_1(|\mathbf{q}|)}{(g_1(|\mathbf{q}|)^2+g_0(|\mathbf{q}|)^2)^{1/2}}\right| \\
\leq&\frac{\alpha}{8\pi^2}\int_B d\mathbf{q}\frac{4|\mathbf{p}||\mathbf{q}|}{(|\mathbf{p}|^2-|\mathbf{q}|^2)^2}\frac{(1+\epsilon)|\mathbf{q}|}{(|\mathbf{q}|^2+m_0^2)^{1/2}} \\
\leq&\frac{8\alpha(1+\epsilon)}{9\pi^2}|\mathbf{p}|\int_B d\mathbf{q}\frac{1}{|\mathbf{q}|^2}\frac{1}{(|\mathbf{q}|^2+m_0^2)^{1/2}} \\
\leq&\frac{32\alpha(1+\epsilon)}{9\pi}|\mathbf{p}|\operatorname{arsinh}(\Lambda/m_0).
\end{aligned} \tag{34}
$$

Thus, we obtain the bound

$$
g_1(|\mathbf{p}|) \leq |\mathbf{p}|\{1+\frac{50\alpha}{9\pi}(1+\epsilon)\operatorname{arsinh}(\Lambda/m_0)\}. \tag{35}
$$

□

Theorem 1. *If $Y := \alpha\operatorname{arsinh}(\Lambda/m_0)/\pi < 9/50$, if $(2+\epsilon+\delta)Y < 1$ and if ϵ,δ satisfy the conditions of Lemma 1 then T is a contraction.*

Proof. Thanks to Lemma 1 we only need to establish the contraction property. We first note that for positive real numbers $x, y, \tilde{x}, \tilde{y}$

$$
\left|\frac{x}{(x^2+y^2)^{1/2}} - \frac{\tilde{x}}{(\tilde{x}^2+\tilde{y}^2)^{1/2}}\right| \leq \frac{|\eta|}{\xi^2+\eta^2}((x-\tilde{x})^2+(y-\tilde{y})^2)^{1/2} \tag{36}
$$

where (ξ,η) is some point on the line between (x,y) and $(\tilde{x},\tilde{y})$.

Thus we get

$$
\begin{aligned}
&|[T_0(g)-T_0(\tilde g)](u)|+|[T_1(g)-T_1(\tilde g)](u)|\\
\le&\frac{\alpha}{2\pi}\int_0^\Lambda dv\frac{v}{u}\left[Q_0\left(\frac12\left(\frac{u}{v}+\frac{v}{u}\right)\right)\left|\frac{g_0(v)}{\sqrt{g_0(v)^2+g_1(v)^2}}-\frac{\tilde g_0(v)}{\sqrt{\tilde g_0(v)^2+\tilde g_1(v)^2}}\right|\right.\\
&\qquad\left.+Q_1\left(\frac12\left(\frac{u}{v}+\frac{v}{u}\right)\right)\left|\frac{g_1(v)}{\sqrt{g_0(v)^2+g_1(v)^2}}-\frac{\tilde g_1(v)}{\sqrt{\tilde g_0(v)^2+\tilde g_1(v)^2}}\right|\right]\\
\le&\frac{\alpha}{2\pi}\int_0^\Lambda dv\frac{v}{u}\left[Q_0\left(\frac12\left(\frac{u}{v}+\frac{v}{u}\right)\right)\frac{(1+\epsilon)v}{v^2+m_0^2}+Q_1\left(\frac12\left(\frac{u}{v}+\frac{v}{u}\right)\right)\frac{(1+\delta)m_0}{v^2+m_0^2}\right]\\
&\qquad\times\left[(g_0(v)-\tilde g_0(v))^2+(g_1(v)-\tilde g_1(v))^2\right]^{1/2}\\
\le&\frac{\alpha}{2\pi}\int_0^\Lambda dv\frac{v}{u}\left[Q_0\left(\frac12\left(\frac{u}{v}+\frac{v}{u}\right)\right)\frac{(1+\epsilon)v}{v^2+m_0^2}\right.\\
&\qquad\left.+Q_1\left(\frac12\left(\frac{u}{v}+\frac{v}{u}\right)\right)\frac{(1+\delta)m_0}{v^2+m_0^2}\right]\|\mathbf{g}-\tilde{\mathbf{g}}\|\\
\le&\frac{\alpha}{2\pi}\int_0^\Lambda dv\frac{v}{u}Q_0\left(\frac12\left(\frac{u}{v}+\frac{v}{u}\right)\right)\frac{(1+\epsilon)v+(1+\delta)m_0}{v^2+m_0^2}\|\mathbf{g}-\tilde{\mathbf{g}}\|\\
\le&(2+\epsilon+\delta)Y\|\mathbf{g}-\tilde{\mathbf{g}}\|
\end{aligned}
\tag{37}
$$

where $Y=\alpha\,\mathrm{arsinh}(\Lambda/m_0)/\pi$ as in Lemma 1. In the last line we have simply noted that the integrals are smaller than the corresponding integral in (32). □

Corollary 1. *The map T has a unique fixed point, if $Y<1/7$.*

Proof. Take $\epsilon=50Y/(9-50Y)$ and $\delta=Y/(1-Y)$ so that T maps $S_{\epsilon,\delta}$ into itself. Then, if $Y<Y_0$, the contraction condition will be satisfied. □

For $\alpha=1/137$, i.e., the physical value of the fine structure constant – the condition $Y<1/7$ is fulfilled for $\Lambda/m\le e^{18}$.

4. Properties of the Renormalized Hamiltonian

What we have shown up to now is that the bare Hamiltonian in (4) is equivalent, apart from some additive constants, to a renormalized Hamiltonian. This renormalized Hamiltonian has the form (6), but the quadratic term is only approximately the one given in (7). What takes the place of $D_{Z,m}$ is given in (17) and (19). We see immediately from (22), (23) that the solution is a continuous pair of functions. We also see that as long as m_0 is not zero the functions g_0 and g_1 are positive and behave properly for small $|\mathbf{p}|$, i.e., g_0 is constant and g_1 is proportional to $|\mathbf{p}|$. To relate this to $D_{Z,m}$ we first factor out the constant $\lim_{u\to0+}(g_1(u)/u)$ and call this Z^{-1}. As is evident from (21), this constant $1/Z$ is larger than one — as it should be.

The next thing to verify is that $Zg_0(0)$ is bigger than m_0, since this is the renormalized mass m that appears in (7). In other words, we have to verify that

$$\frac{m}{m_0}Z^{-1}=\frac{g_0(0)}{m_0}>\lim_{\mathbf{p}\to0}\frac{g_1(|\mathbf{p}|)}{|\mathbf{p}|}=Z^{-1}. \tag{38}$$

We shall, in fact, prove more than this:

Theorem 2. *Assuming that $Y < 1/7$, the unique solution to the equation* (20) *and* (21) *mentioned has the property*

$$\frac{g_0(|\mathbf{p}|)}{m_0} > \frac{g_1(|\mathbf{p}|)}{|\mathbf{p}|} \tag{39}$$

for all $\mathbf{p} \neq \mathbf{0}$.

Proof. Define $\tilde{S}_{\delta,\epsilon}$ to be the subset of $S_{\delta,\epsilon}$ on which

$$\frac{g_0(|\mathbf{p}|)}{m_0} \geq \frac{g_1(|\mathbf{p}|)}{|\mathbf{p}|} \tag{40}$$

for all $\mathbf{p} \neq \mathbf{0}$. If we can show that T also leaves $\tilde{S}_{\delta,\epsilon}$ invariant, then we have shown the wanted inequality (39) on the solution in $S_{\delta,\epsilon}$ (by the uniqueness of the solution on $S_{\delta,\epsilon}$) except for the possibility that equality also can occur. This is so, since we can apply the fixed point argument not only to $S_{\delta,\epsilon}$ but also to $\tilde{S}_{\delta,\epsilon}$.

Showing that (40) holds is – according to (22) and (23) – equivalent to showing that

$$\begin{aligned} \int_0^\Lambda dv \frac{v}{u} Q_0\left(\frac{1}{2}\left(\frac{u}{v}+\frac{v}{u}\right)\right) \frac{g_0(v)/m_0}{(g_1(v)^2+g_0(v)^2)^{1/2}} \\ > \int_0^\Lambda dv \frac{v}{u} Q_1\left(\frac{1}{2}\left(\frac{u}{v}+\frac{v}{u}\right)\right) \frac{g_1(v)/u}{(g_1(v)^2+g_0(v)^2)^{1/2}} \end{aligned} \tag{41}$$

holds. Now, using the Inequality (40) and the fact that the factors with the roots are monotone functions g_0 and g_1 respectively, it is enough to show that

$$\begin{aligned} \int_0^\Lambda dv \frac{v}{u} Q_0\left(\frac{1}{2}\left(\frac{u}{v}+\frac{v}{u}\right)\right) \frac{1}{(m_0^2+v^2)^{1/2}} \\ > \int_0^\Lambda dv \frac{v}{u} Q_1\left(\frac{1}{2}\left(\frac{u}{v}+\frac{v}{u}\right)\right) \frac{v/u}{(m_0^2+v^2)^{1/2}} \end{aligned} \tag{42}$$

holds. To proceed, we now compare the integrands pointwise. Using the explicit expressions (24) and (25) for Q_0 and Q_1, and with $t = v/u$, we will have shown that the integrand of the left hand side pointwise majorizes the one on the right hand side of (42) if

$$\log\left|\frac{t+1}{t-1}\right| < \frac{2t}{|t^2-1|} \tag{43}$$

holds for all $t \neq 1$. By symmetry, we only have to consider (43) for $t > 1$. To this end we exponentiate (43)

$$\frac{t+1}{t-1} < \exp\left(\frac{2t}{t^2-1}\right)$$

and expand the exponential up to second order. (Note that this gives a lower bound on the exponential because the argument of the exponential function is positive.) I.e., it suffices to show that

$$\frac{t+1}{t-1} < 1 + \frac{2t}{t^2-1} + 2\frac{t^2}{(t^2-1)^2} \tag{44}$$

which follows by direct computation.

Having established (2) we see that (44) indeed gives strict inequality in (2) for the unique fixed point in $\tilde{S}_{\delta,\epsilon}$. □

5. Interpretations of our Results

What we have done is start with the 'bare' Hamiltonian $\mathbb{H}^{\text{bare}}$ in (4), in which the interaction is the closest analog to the classical electrostatic energy of a field. After some analysis, we found in section 3 that Z and m could be uniquely chosen so that for momenta much less than the ultraviolet cutoff Λ, $\mathbb{H}^{\text{bare}} = \mathbb{H}_\alpha^{\text{ren}}$ plus a well defined infinite constant. Of course the field Ψ in the two cases is different. *Formally*, they differ by a Bogolubov transformation, but in fact an inequivalent, α *dependent*, representation of the CAR is needed. In section 4 we showed that not only is $Z^{-1}m > m_0$ but more is true, namely $m > m_0$, and this is comforting physically.

We emphasize that we have not 'integrated out' any field variables. Our new Hamiltonian $\mathbb{H}_\alpha^{\text{ren}}$ is the same as the original one — on a purely formal level. It is suggestive, nevertheless, that a good deal of the electrostatic energy has now been incorporated into the leading term $\int d^3x : \Psi^*(x) D_{Z,m}\Psi(x) :$ and that the remaining interaction is somehow less important than the original one. It is, after all, normal ordered, which means it vanishes on one-electron states, if we define such states by $\Psi^*|0\rangle$, where $|0\rangle$ is the vacuum of the *new* Ψ. While this makes sense perturbatively, it is, however, misleading because the new vacuum (the state of lowest energy) is surely not the obvious choice $|0\rangle$.

If we drop the new interaction term $\mathbb{W}_\alpha^{\text{ren}}$ we are left with $\int d^3x : \Psi^*(x) D_{Z,m}\Psi(x) :$ as our Hamiltonian. Unfortunately this is not the Hamiltonian of a Dirac operator (even at low momenta) because of the factor Z^{-1}. We have called this a wave function renormalization, but that is not really in the spirit of renormalizing the one-electron states (which is what is usually done) and is, instead, a renormalization of an operator. One school, (Källén [5]) thinks it is proper to speak of a renormalized Ψ by incorporating a factor $Z^{-1/2}$ into Ψ, but this changes the anti-commutation relations! Note that this formulation requires renormalizing the bare mass m_0 to $g_0(\mathbf{0})/Z$ and renormalizing the bare charge e to e/Z.

Another point of view is to regard $Z^{-1}m$ as the physical mass and to change $Z^{-1}\mathbf{p}$ into $\mathbf{p}$ by a unitary transformation (which is nothing other than a change of length scale). This has the disadvantage of changing the speed of light or Planck's constant. It also would mean a different scale change for particles of different mass, e.g., muons. As opposed to Källén's procedure, this requires renormalizing the mass from m_0 to $g_0(\mathbf{0})$ only, but there is no need for charge renormalization.

Another possibility is to bring out the factor Z^{-1} as a multiplier of the whole Hamiltonian, which would mean changing the fine structure constant to αZ, i.e., a *charge* renormalization but now from e to $e/Z^{1/2}$ only. The obvious problem here is that since the Hamiltonian is the generator of time translation, this means a change of the time scale (which, again, depends on the particle in question).

Doubtless, different people will have different opinions about these matters. We do not wish to commit ourselves to any point of view. But it is our opinion that the construction of $\mathbb{H}_\alpha^{\text{ren}}$ is a significant piece of the puzzle of constructing a non-perturbative QED.

Acknowledgement: The authors thank Dirk Hundertmark for a useful discussion. Financial support of the European Union, TMR grant FMRX-CT 96-0001, the U.S. National Science Foundation, grant PHY98-20650, and NATO, grant CRG96011 is acknowledged.

References

[1] Volker Bach, Jean-Marie Barbaroux, Bernard Helffer, and Heinz Siedentop. Stability of matter for the Hartree-Fock functional of the relativistic electron-positron field. *Doc. Math.*, 3:353–364 (electronic), 1998.

[2] Volker Bach, Jean-Marie Barbaroux, Bernard Helffer, and Heinz Siedentop. On the stability of the relativistic electron-positron field. *Commun. Math. Phys.*, 201:445–460, 1999.

[3] James D. Drell and Sidney D. Drell. *Relativistic Quantum Fields.* International Series in Pure and Applied Physics. McGraw-Hill, New York, 1 edition, 1965.

[4] Bernard Helffer and Heinz Siedentop. Form perturbations of the second quantized Dirac field. *Math. Phys. Electron. J.*, 4:Paper 4, 16 pp. (electronic), 1998.

[5] G. Källén. Quantenelektrodynamik. In S. Flügge, editor, *Prinzipien der Quantentheorie I*, volume V/1 of *Handbuch der Physik*, pages 169–364. Springer-Verlag, Berlin, 1 edition, 1958.

[6] Bernd Thaller. *The Dirac Equation.* Texts and Monographs in Physics. Springer-Verlag, Berlin, 1 edition, 1992.

Departments of Mathematics and Physics, Jadwin Hall, Princeton University, P.O.B. 708, Princeton, NJ 08544-0708, USA
E-mail address: `lieb@princeton.edu`

Mathematik I, Universität Regensburg, 93040 Regensburg, Germany
E-mail address: `heinz.siedentop@mathematik.uni-regensburg.de`