hep-th/9705006 v4   30 Mar 1999

Definition of the Dirac Sea in the Presence of External Fields

Felix Finster [1]

Mathematics Department
Harvard University

Abstract

It is shown that the Dirac sea can be uniquely defined for the Dirac equation with general interaction, if we impose a causality condition on the Dirac sea. We derive an explicit formula for the Dirac sea in terms of a power series in the bosonic potentials.

The construction is extended to systems of Dirac seas. If the system contains chiral fermions, the causality condition yields a restriction for the bosonic potentials.

1   Introduction

The Dirac equation has solutions of negative energy, which have no meaningful physical interpretation. This popular problem of relativistic quantum mechanics was originally solved by Dirac's concept that all negative-energy states are occupied in the vacuum forming the so-called Dirac sea. Fermions and anti-fermions are then described by positive-energy states and "holes" in the Dirac sea, respectively. Although this vivid picture of a sea of interacting particles is nowadays often considered not to be taken too literally, the construction of the Dirac sea also plays a crucial role in quantum field theory. There it corresponds to the formal exchanging of creation and annihilation operators for the negative-energy states of the free field theory.

Usually, the Dirac sea is only constructed in the vacuum. This is often considered to be sufficient, because the interacting system can be described by a perturbation of the vacuum.

Unfortunately, the situation is more difficult: In relativistic quantum mechanics with interaction, the fermionic wave functions are solutions of the Dirac equation

$$(i\partial\!\!\!/ + \mathcal{B} - m)\,\tilde{\Psi} \;=\; 0, \tag{1.1}$$

where the operator $\mathcal{B}$ is composed of the bosonic potentials (for example, we can describe the electromagnetic interaction by choosing $\mathcal{B} = e A\!\!\!/$ with the electromagnetic potential A).

In contrast to the free Dirac equation $(i\partial\!\!\!/ - m)\,\Psi = 0$, it is not obvious how to characterize the negative-energy solutions of (1.1). Qualitatively, the problem is that the perturbation $\mathcal{B}$ leads to a mixing of the free solutions and destroys the natural splitting into solutions of positive and negative energy. As a consequence, it is not clear how the Dirac sea of the system (1.1) can be constructed. We point out that this problem is not solved by a simple perturbation expansion in $\mathcal{B}$; it is then hidden in the non-uniqueness of this expansion (see section 2 for details). In quantum field theory, the problem of defining the Dirac sea is even more complicated, because the virtual pair creation/annihilation must be taken into account. We will not deal these problems here and restrict to the limit of "classical" potentials and wave functions. Nevertheless, our considerations are also relevant for quantum field theory, because it is in many situations (e.g. for a quantum system in a classical background field) preferable to use the Dirac equation (1.1) as the starting point for the fermionic field quantization. In this sense, the construction of the Dirac sea of (1.1) is preliminary for the description of interacting quantum fields.

We conclude that the definition of the Dirac sea is basic for a reasonable physical interpretation of the Dirac equation (1.1). In the present paper, we will discuss the difficulty in constructing the Dirac sea and finally solve the problem in terms of a formal perturbation expansion in $\mathcal{B}$. Before starting the analysis, we describe the problem in more mathematical terms: Every solution of the free Dirac equation $(i\partial\!\!\!/ - m)\,\Psi = 0$ is a linear combination of plane wave solutions of the form

$$\Psi(t, \vec{x}) \;=\; e^{-i(\omega t - \vec{k}\vec{x})}\,\chi_{\omega,\vec{k}} \quad , \qquad \omega \;=\; \pm\sqrt{\vec{k}^2 + m^2}$$

with a 4-spinor $\chi_{\omega,\vec{k}}$ which is independent of t and $\vec{x}$. The sign of ω gives a natural splitting of the solutions into solutions of positive and negative frequency. Identifying frequency and energy via Planck's formula, these solutions are commonly called the positive and negative energy solutions of the free Dirac equation. Since the simple identification of frequency and energy

might lead to confusion (sometimes the "energy" of a negative-frequency state denotes the positive energy of the corresponding anti-particle state), we prefer the notion of positive and negative "frequency" in the following. We denote the negative-frequency solutions by $\Psi_{\vec{k}a}$, where $\vec{k}$ is the momentum and $a = 1, 2$ are the two spin states (for an explicit formula for $\Psi_{\vec{k}a}$ see e.g. [1]). If the states $\Psi_{\vec{k}a}$ were normalized with respect to the usual scalar product

$$(\Psi \mid \Phi) \;=\; \int_{\mathrm{IR}^3} (\overline{\Psi} \, \gamma^0 \, \Phi)(t, \vec{x}) \, d\vec{x} \;\;, \qquad \overline{\Psi} = \Psi^* \gamma^0 \qquad , \qquad (1.2)$$

we could form the projector $P_{<\Psi_{\vec{k}a}>}$ on the one-dimensional subspace $<\Psi_{\vec{k}a}>$ by

$$\left(P_{<\Psi_{\vec{k}a}>} \Psi \right)(t, \vec{x}) \;=\; \int_{\mathrm{IR}^3} \left(\Psi_{\vec{k}a}(t, \vec{x}) \, \overline{\Psi_{\vec{k}a}(t, \vec{y})} \right) \gamma^0 \, \Psi(t, \vec{y}) \, d\vec{y} \qquad .$$

In this sense, the product $\Psi_{\vec{k}a}(x) \, \overline{\Psi_{\vec{k}a}(y)}$ would be the kernel of the projector on $<\Psi_{\vec{k}a}>$, and the sum over all negative-frequency states would yield the projector on the whole Dirac sea. Unfortunately, the wave functions $\Psi_{\vec{k}a}$ are not normalizable. We could arrange normalizable states by considering the system in finite three-volume, but we do not want to do this here. It is more appropriate for our purpose to formally build up a projector on all negative-frequency states by integrating over the momentum parameter

$$P(x, y) \;=\; \sum_{a=1,2} \int_{\mathrm{IR}^3} \Psi_{\vec{k}a}(x) \, \overline{\Psi_{\vec{k}a}(y)} \, d\vec{k} \qquad , \qquad (1.3)$$

which can be rewritten as the integral over the lower mass shell

$$=\; \int_{\mathrm{IR}^3} \frac{d^4k}{(2\pi)^4} \, (\slashed{k} + m) \, \delta(k^2 - m^2) \, \Theta(-k^0) \, e^{-ik(x-y)}$$

(Θ denotes the Heavyside function $\Theta(x) = 1$ for $x \geq 0$ and $\Theta(x) = 0$ otherwise). $P(x, y)$ is a well-defined tempered distribution which solves the free Dirac equation $(i \slashed{\partial}_x - m) \, P(x, y) = 0$. We can use it to characterize the Dirac sea in the vacuum. Our aim is to introduce a corresponding distribution $\tilde{P}$ for the Dirac equation with interaction (1.1). The construction of $\tilde{P}$ must be unique in a sense which we will discuss and specify later. We will assume the perturbation $\mathcal{B}$ to be a differential operator on the wave functions. Furthermore, it shall be Hermitian with respect to the (indefinite) scalar product

$$<\Psi \mid \Phi> \;=\; \int \overline{\Psi(x)} \, \Phi(x) \, d^4x \qquad . \qquad (1.4)$$

For an electromagnetic potential $\mathcal{B} = e\slashed{A}$, these assumptions are satisfied because $\slashed{A} = \gamma^0 \slashed{A}^\dagger \gamma^0$. In addition, $\mathcal{B}$ can be composed of the scalar, pseudoscalar, pseudovector and bilinear potentials as e.g. discussed in [2]. According to [3], $\mathcal{B}$ also allows for the description of the gravitational field.

2   Non-Uniqueness of the Simple Perturbation Expansion

Our first idea for the construction of $\tilde{P}$ is to calculate solutions $\tilde{\Psi}_{\vec{k}a}$ of (1.1) with a perturbation expansion in $\mathcal{B}$ and to define $\tilde{P}$ in analogy to (1.3) by

$$\tilde{P}(x,y) \;=\; \sum_{a=1,2} \int_{\mathbb{R}^3} \tilde{\Psi}_{\vec{k}a}(x)\,\overline{\tilde{\Psi}_{\vec{k}a}(y)}\,d\vec{k} \qquad . \tag{2.1}$$

We start with a discussion of this method in a perturbation calculation to first order. This is quite elementary and will nevertheless explain the basic difficulty. For the perturbation calculation, we need a Green's function $s(x,y)$ of the free Dirac operator, which is characterized by the distributional equation

$$(i\partial\!\!\!/_x - m)\,s(x,y) \;=\; \delta^4(x-y) \qquad . \tag{2.2}$$

To first order, the perturbed eigenstates $\tilde{\Psi}_{\vec{k}a}$ are then given by

$$\tilde{\Psi}_{\vec{k}a}(x) \;=\; \Psi_{\vec{k}a}(x) \;-\; \int d^4y\; s(x,y)\,\mathcal{B}_y\,\Psi_{\vec{k}a}(y) \;+\; \mathcal{O}(\mathcal{B}^2) \qquad , \tag{2.3}$$

as can be verified by substituting into (1.1). We insert this formula into (2.1) and obtain

$$\tilde{P}(x,y) \;=\; P(x,y) - \int d^4z\; [s(x,z)\,\mathcal{B}_z\,P(z,y) \;+\; P(x,z)\,\mathcal{B}_z\,s^*(z,y)] + \mathcal{O}(\mathcal{B}^2), \tag{2.4}$$

where we used that $\mathcal{B}$ is Hermitian with respect to the scalar product (1.4), and where $s^*(z,y)$ is given by $s^*(z,y) = \gamma^0\,s(y,z)^\dagger\,\gamma^0$. It is convenient to view the distributions $s(x,y), P(x,y)$ as integral kernels of corresponding operators s, P. Then we can write (2.4) with operator products

$$\tilde{P} \;=\; P \;-\; s\,\mathcal{B}\,P \;-\; P\,\mathcal{B}\,s^* \;+\; \mathcal{O}(\mathcal{B}^2) \qquad , \tag{2.5}$$

where the superscript '*' denotes the adjoint with respect to the scalar product (1.4).

Equation (2.5) gives a possible definition for $\tilde{P}$. As apparent problem, the construction depends on the choice of the Green's function. For example, we could have chosen for s either the advanced or the retarded Green's function $s_m^\vee, s_m^\wedge$, which are in momentum space as usual given by

$$s_m^\vee(k) \;=\; \lim_{0<\varepsilon\to0} \frac{k\!\!\!/ + m}{k^2 - m^2 - i\varepsilon k^0} \quad , \quad s_m^\wedge(k) \;=\; \lim_{0<\varepsilon\to0} \frac{k\!\!\!/ + m}{k^2 - m^2 + i\varepsilon k^0} \tag{2.6}$$

More systematically, the arbitrariness of our construction is described as follows: According to (2.2), the difference between two Green's functions is a solution of the free Dirac equation. We can thus represent s in the form

$$s(x,y) \;=\; s_m^\vee(x,y) \;+\; a(x,y) \qquad ,$$

where $a(x, y)$ is in the x-variable a linear combination of the plane-wave solutions, i.e.

$$a(x, y) = \sum_{a=1}^{4} \int_{\mathbb{R}^3} \Psi_{\vec{k}a}(x) \, c_{\vec{k}a}(y) \, d\vec{k}$$

with (generally complex) functions $c_{\vec{k}a}(y)$, where $\Psi_{\vec{k}a}, a = 3, 4$ denote the plane-wave solutions of positive frequency. We substitute into (2.5) and obtain

$$\tilde{P} = P - s_m^{\vee} \mathcal{B} P - P \mathcal{B} s_m^{\wedge} - (a \mathcal{B} P + P \mathcal{B} a^*) + \mathcal{O}(\mathcal{B}^2) \qquad . \quad (2.7)$$

The expression in the brackets maps solutions of the free Dirac equation into each other and vanishes otherwise. We can thus write it in the form

$$(a \mathcal{B} P + P \mathcal{B} a^*)(x, y) = \sum_{a,b=1}^{4} \int_{\mathbb{R}^3} d\vec{k}_1 \int_{\mathbb{R}^3} d\vec{k}_2 \, \Psi_{\vec{k}_1 a}(x) \, g_{ab}(\vec{k}_1, \vec{k}_2) \, \overline{\Psi_{\vec{k}_2 b}(y)}$$

$$(2.8)$$

with suitable functions $g_{ab}(\vec{k}_1, \vec{k}_2)$. This representation of $\tilde{P}$ can also be understood directly: The contribution (2.8) describes a mixing of the solutions $\Psi_{\vec{k}a}$ of the free Dirac equation. To the considered first order in $\mathcal{B}$, it vanishes in the Dirac equation $(i \partial\!\!\!/ + \mathcal{B} - m) \, \tilde{P} = 0$. Thus we cannot fix this contribution with the Dirac equation, it remains undetermined in our method. According to (2.7), this is the only arbitrariness of the construction; the other contributions to $\tilde{P}$ are unique.

In higher order perturbation theory, the non-uniqueness can be understood similarly, although the situation is more complicated: For a given Green's function s, we can construct a solution $\tilde{\Psi}_{\vec{k}a}$ of the Dirac equation (1.1) by the formal perturbation series

$$\tilde{\Psi}_{\vec{k}a} = \sum_{n=0}^{\infty} (-s \, \mathcal{B})^n \, \Psi_{\vec{k}a} \qquad , \qquad (2.9)$$

as is verified by substituting into (1.1). Actually, this is a very special ansatz. For example, we can use different Green's functions in every order of the perturbation calculation, which leads to the more general formula

$$\tilde{\Psi}_{\vec{k}a} = \Psi_{\vec{k}a} + \sum_{n=1}^{\infty} (-1)^n \, s^{(n)} \, \mathcal{B} \, \cdots \, s^{(2)} \, \mathcal{B} \, s^{(1)} \, \mathcal{B} \, \Psi_{\vec{k}a} \qquad (2.10)$$

with a whole series of arbitrary Green's functions $s^{(1)}$, $s^{(2)}$, etc.. Once we have a formula for $\tilde{\Psi}_{\vec{k}a}$, the non-uniqueness of $\tilde{P}$ can again be discussed by substituting into (2.1). In generalization of (2.8), the arbitrariness of the construction is described by a contribution to $\tilde{P}(x, y)$ of the form

$$\sum_{a,b=1}^{4} \int_{\mathbb{R}^3} d\vec{k}_1 \int_{\mathbb{R}^3} d\vec{k}_2 \, \tilde{\Psi}_{\vec{k}_1 a}(x) \, g_{ab}(\vec{k}_1, \vec{k}_2) \, \overline{\tilde{\Psi}_{\vec{k}_2 b}(y)} \qquad ,$$

which mixes perturbed eigenstates $\tilde{\Psi}_{\vec{k}a}$ and vanishes in the Dirac equation $(i\partial\!\!\!/ + \mathcal{B} - m)\,\tilde{P} = 0$. The dependence of $g_{ab}(\vec{k}_1, \vec{k}_2)$ on $\mathcal{B}$ and on the Green's functions $s^{(n)}$ is rather involved, however, and we need not go into the details here.

To summarize, a simple perturbation expansion in $\mathcal{B}$ is not unique and therefore does not allow a meaningful definition of $\tilde{P}$. In the ansatz (2.10), for example, we should find a way to specify the Green's functions $s^{(n)}$. This cannot be done with the Dirac equation (1.1), and we must therefore look for additional input to completely determine $\tilde{P}$. Our basic idea is to apply some causality principle. For example, it might seem a reasonable condition to impose that $\tilde{P}(x, y)$ only depends on $\mathcal{B}$ in the "diamond" $(L_x^\vee \cap L_y^\wedge) \cup (L_x^\wedge \cap L_y^\vee)$, where

$$L_x^\vee \;=\; \left\{ y \mid (y - x)^2 \geq 0,\; y^0 - x^0 \geq 0 \right\} \tag{2.11}$$

$$L_x^\wedge \;=\; \left\{ y \mid (y - x)^2 \geq 0,\; y^0 - x^0 \leq 0 \right\} \tag{2.12}$$

denote the future and past light cones around x, respectively. If we want to study conditions of this type, it is no longer useful to look at the perturbation expansion for the individual states $\Psi_{\vec{k}a}(x)$ (because these states only depend on one argument x). We must take into account for the perturbation expansion that P is composed of many states in a specific way.

3   The Causal Perturbation Expansion

In preparation, we first describe how the perturbation expansion for the advanced and retarded Green's functions can be performed uniquely: The support of the distribution $s_m^\vee(x, y)$ is in the future light cone $y \in L_x^\vee$ (this can be checked by calculating the Fourier transform of (2.6) with contour integrals). As a consequence, the perturbation operator $\mathcal{B}(z)$ only enters into the operator product

$$(s_m^\vee\,\mathcal{B}\,s_m^\vee)(x, y) \;=\; \int d^4z\; s_m^\vee(x, z)\,\mathcal{B}(z)\,s_m^\vee(z, y) \tag{3.1}$$

for $z \in L_x^\vee \cap L_y^\wedge$. In this sense, the expression (3.1) is *causal*. Especially, the support of (3.1) is again in the future light cone. It follows by iteration that the higher powers

$$s_m^\vee\,\mathcal{B}\,s_m^\vee\,\mathcal{B}\,\cdots\mathcal{B}\,s_m^\vee\,\mathcal{B}\,s_m^\vee$$

are also causal and have their support in the upper light cone. We define the perturbed advanced Green's function as the formal sum over these operator products,

$$\tilde{s}_m^\vee \;=\; \sum_{k=0}^\infty \left(-s_m^\vee\,\mathcal{B}\right)^k s_m^\vee \qquad . \tag{3.2}$$

Accordingly, the perturbed retarded Green's function is defined by

$$\tilde{s}_m^\wedge \;=\; \sum_{k=0}^{\infty} \left(-s_m^\wedge\, \mathcal{B}\right)^k s_m^\wedge \qquad . \tag{3.3}$$

These operators satisfy the defining equations for the perturbed Green's functions

$$(i\partial\!\!\!/ - m + \mathcal{B})\, \tilde{s}_m^\vee \;=\; \mathbb{1} \;=\; (i\partial\!\!\!/ - m + \mathcal{B})\, \tilde{s}_m^\wedge \qquad , \tag{3.4}$$

as is verified directly.

Notice that the perturbation expansion for the Green's functions becomes unique by the condition that the contribution to $\tilde{s}_m^\vee, \tilde{s}_m^\wedge$ to every order has its support in the future and past light cones, respectively. We want to take this construction as the guiding line for the perturbation expansion of P. Unfortunately, the method cannot be directly applied to the Dirac sea, because the distribution $P(x,y)$ does not vanish for space-like $y-x$, and we thus have no notion of causality. As way out, we decompose the free Dirac sea in the form

$$P(x,y) \;=\; \frac{1}{2}\left(p_m(x,y)\; -\; k_m(x,y)\right) \tag{3.5}$$

with the tempered distributions

$$p_m(x,y) \;=\; \int \frac{d^4k}{(2\pi)^4}\,(k\!\!\!/ + m)\,\delta(k^2 - m^2)\, e^{-ik(x-y)} \tag{3.6}$$

$$k_m(x,y) \;=\; \int \frac{d^4k}{(2\pi)^4}\,(k\!\!\!/ + m)\,\delta(k^2 - m^2)\,\epsilon(k^0)\, e^{-ik(x-y)} \tag{3.7}$$

(ϵ denotes the step function $\epsilon(x) = 1$ for $x \geq 0$ and $\epsilon(x) = -1$ otherwise). We also consider $p_m(x,y)$ and $k_m(x,y)$ as integral kernels of corresponding operators p_m, k_m. The operator p_m is built up as a formal sum over the projectors on all solutions of the Dirac equation and can be viewed as a spectral projector of the free Dirac operator. The definition of k_m differs from p_m by a relative minus sign for the states on the upper and lower mass shell. As a consequence of this relative minus sign, the Fourier integral (3.7) vanishes if $y-x$ is space-like (this can be seen from Lorentzian invariance and a symmetry argument for $k = (0, \vec{k})$). Thus $k_m(x,y)$ is causal in the sense that it has the support in the light cone $y \in L_x^\vee \cup L_x^\wedge$. This makes it possible to uniquely express its perturbation expansion in terms of the perturbed Green's functions: We substitute the distributional equation

$$\lim_{0 < \varepsilon \to 0} \left(\frac{1}{x - i\varepsilon} - \frac{1}{x + i\varepsilon}\right) \;=\; 2\pi i\,\delta(x)$$

into the formula for k_m in momentum space,

$$k_m(p) \;=\; (p\!\!\!/ + m)\,\delta(p^2 - m^2)\,\epsilon(p^0)$$

$$= \frac{1}{2\pi i} (\not{p} + m) \lim_{0 < \varepsilon \to 0} \left[\frac{1}{p^2 - m^2 - i\varepsilon} - \frac{1}{p^2 - m^2 + i\varepsilon} \right] \epsilon(p^0)$$

$$= \frac{1}{2\pi i} (\not{p} + m) \lim_{0 < \varepsilon \to 0} \left[\frac{1}{p^2 - m^2 - i\varepsilon p^0} - \frac{1}{p^2 - m^2 + i\varepsilon p^0} \right] \quad ,$$

and obtain with (2.6) a simple relation between k_m and $s_m^\vee$, $s_m^\wedge$,

$$k_m = \frac{1}{2\pi i} \left(s_m^\vee - s_m^\wedge \right) \qquad . \tag{3.8}$$

We extend this relation to the case with external fields:

Def. 3.1 *We define the operator $\tilde{k}_m$ by*

$$\tilde{k}_m = \frac{1}{2\pi i} \left(\tilde{s}_m^\vee - \tilde{s}_m^\wedge \right) \tag{3.9}$$

with the Green's functions (3.2),(3.3).

According to (3.4), $\tilde{k}_m$ really is a solution of the Dirac equation $(i\not{\partial} + \mathcal{B} - m)\, \tilde{k}_m = 0$.

In order to explain the significance of this construction, we point out that the factor $\epsilon(k^0)$ in (3.7) describes the splitting of the solutions of the free Dirac equation into solutions of positive and negative frequency. With the introduction of $\tilde{k}_m$, we were able to uniquely generalize this splitting to the case with external fields. This solves the basic problem in defining the Dirac sea. It remains to perform the perturbation expansion for $\tilde{p}_m$. On a formal level, this is very easy, because we can remove the relative minus sign for the positive and negative frequency states by taking the absolute value of $\tilde{k}_m$,

$$\tilde{p}_m \stackrel{\text{formally}}{:=} \sqrt{\tilde{k}_m^2} \qquad . \tag{3.10}$$

This gives a unique definition for $\tilde{p}_m$. Since $\tilde{k}_m$ is composed of eigenstates of the perturbed Dirac operator with eigenvalue m, it follows automatically that $(i\not{\partial} + \mathcal{B} - m)\, \tilde{p}_m = 0$.

Unfortunately, it requires some effort to convert the formal relation (3.10) into a mathematical definition. The problem is that the square $\tilde{k}_m^2$ is ill-defined; furthermore we want to write $\tilde{p}_m$ as a power series in $\mathcal{B}$. These problems are solved in the following theorem. The reader who is not so interested in the technical details and the combinatorics of the expansion may skip the proof. For the statement of the theorem, we need some notation: We work with the Green's function

$$s_m = \frac{1}{2}(s_m^\vee + s_m^\wedge) \qquad , \tag{3.11}$$

which has the advantage of being Hermitian (with respect to the scalar product (1.4)). Furthermore, we introduce the series of operator products

$$b_m^< = \sum_{k=0}^{\infty} (-s_m \, \mathcal{B})^k \quad , \qquad b_m = \sum_{k=0}^{\infty} (-\mathcal{B} \, s_m)^k \, \mathcal{B} \quad , \qquad b_m^> = \sum_{k=0}^{\infty} (-\mathcal{B} \, s_m)^k$$

and set for $Q \subset N$

$$F_m(Q, n) = \begin{cases} p_m & \text{if } n \in Q \\ k_m & \text{if } n \notin Q \end{cases} \quad .$$

Theorem 3.2 *The relations (3.9),(3.10) uniquely determine the perturbation expansion for k_m and p_m. We have the explicit formulas*

$$\tilde{k}_m = \sum_{\beta=0}^{\infty} (-i\pi)^{2\beta} \, b_m^< \, k_m \, (b_m \, k_m)^{2\beta} \, b_m^> \tag{3.12}$$

$$\tilde{p}_m = \sum_{\beta=0}^{\infty} \sum_{\alpha=0}^{\left[\frac{\beta}{2}\right]} c(\alpha, \beta) \, G_m(\alpha, \beta) \tag{3.13}$$

with the coefficients

$$c(0, 0) = 1 \quad , \tag{3.14}$$

$$c(\alpha, \beta) = \sum_{n=\alpha+1}^{\beta} (-1)^{n+1} \, \frac{(2n-3)!!}{n! \, 2^n} \begin{pmatrix} \beta - \alpha - 1 \\ n - \alpha - 1 \end{pmatrix} \text{ for } \beta \geq 1 \tag{3.15}$$

and the operator products

$$G_m(\alpha, \beta) = \sum_{Q \in \mathcal{P}(\beta+1), \ \#Q = 2\alpha+1} (-i\pi)^{2\beta} \, b_m^< \, F_m(Q, 1) \, b_m k_m b_m \, F_m(Q, 2)$$

$$\times \, b_m k_m b_m \cdots b_m k_m b_m \, F_m(Q, \beta + 1) \, b_m^> \quad , \tag{3.16}$$

where $\mathcal{P}(n)$ denotes the set of subsets of $\{1, \ldots, n\}$ (we use the convention $l!! = 1$ for $l \leq 0$).

Proof: Notice that $(i\partial\!\!\!/ + \mathcal{B} - m) \, b_m^< = 0$. Since all operator products in (3.12),(3.16) have a factor $b_m^<$ at the left, the operators $\tilde{p}_m$, $\tilde{k}_m$ are solutions of the Dirac equation

$$(i\partial\!\!\!/ + \mathcal{B} - m) \, \tilde{p}_m = 0 = (i\partial\!\!\!/ + \mathcal{B} - m) \, \tilde{k}_m \quad .$$

Thus the theorem gives a possible perturbation expansion for p_m and k_m. We must verify that the conditions (3.9),(3.10) are satisfied and show uniqueness.

According to (3.8), the advanced and retarded Green's function can be written in the form

$$s_m^\vee \; = \; s_m + i\pi \, k_m \quad , \qquad s_m^\wedge \; = \; s_m - i\pi \, k_m \qquad . \qquad (3.17)$$

We substitute the sums (3.2),(3.3) into (3.9),

$$\tilde{k}_m \; = \; \frac{1}{2\pi i} \sum_{k=0}^{\infty} \left((-s_m^\vee \, \mathcal{B})^k \, s_m^\vee - (-s_m^\wedge \, \mathcal{B})^k \, s_m^\wedge \right) \qquad , \qquad (3.18)$$

insert (3.17) and expand. This gives a sum of operator products of the form

$$C_1 \, \mathcal{B} \, C_2 \, \mathcal{B} \, \cdots \, \mathcal{B} \, C_{l+1} \qquad \text{with} \qquad C_j = k_m \text{ or } C_j = s_m \qquad .$$

The contributions with an even number of factors k_m have the same sign for the advanced and retarded Green's function and cancel in (3.18). The contributions with an odd number of k_m's occur in every Green's function exactly once and have opposite sign. Using the notation

$$C_m(Q,n) \; = \; \begin{cases} k_m & \text{if } n \in Q \\ s_m & \text{if } n \notin Q \end{cases} \quad , \qquad Q \subset \mathbb{N} \qquad ,$$

we can thus rewrite (3.18) in the form

$$\tilde{k}_m \; = \; \sum_{l=0}^{\infty} (-1)^l \sum_{Q \in \mathcal{P}(l+1), \ \#Q \text{ odd}} (i\pi)^{\#Q-1}$$
$$\times \, C_m(Q,1) \, \mathcal{B} \, C_m(Q,2) \, \mathcal{B} \cdots \mathcal{B} \, C_m(Q,l) \, \mathcal{B} \, C_m(Q,l+1) \qquad .$$

After reordering the sums, this coincides with (3.12).

Next we want to give the relation (3.10) a mathematical sense. For this, we consider $m \geq 0$ as a variable mass parameter. Then we can form products of the operators p_m, k_m by manipulating the arguments of the distributions in momentum space. For example, we have with (3.6)

$$\begin{aligned} p_m(k) \, p_{m'}(k) & = (k\!\!\!/ + m) \, \delta(k^2 - m^2) \, (k\!\!\!/ + m') \, \delta(k^2 - (m')^2) \\ & = (k^2 + (m+m')k\!\!\!/ + mm') \, \delta(m^2 - (m')^2) \, \delta(k^2 - m^2) \\ & = (k^2 + (m+m')k\!\!\!/ + mm') \, \frac{1}{2m} \, \delta(m - m') \, \delta(k^2 - m^2) \\ & = \delta(m - m') \, p_m(k) \qquad , \qquad (3.19) \end{aligned}$$

and similarly with (3.7),

$$p_m \, k_{m'} \; = \; k_{m'} \, p_m \; = \; \delta(m - m') \, k_m \qquad\qquad (3.20)$$
$$k_m \, k_{m'} \; = \; \delta(m - m') \, p_m \qquad . \qquad\qquad (3.21)$$

This formalism has some similarity with the bra/ket notation in quantum mechanics, if the position variable $\vec{x}$ is replaced by the mass parameter m.

Equation (3.19) can be understood directly from the fact that p_m are the spectral projectors of the free Dirac operator; the relations (3.20),(3.21) reflect the relative minus sign in k_m for the states on the upper and lower mass shell. Especially one sees that $k_m\, k_{m'} = p_m\, p_{m'}$. This relation can be extended to the case with interaction,

$$\tilde{p}_m\, \tilde{p}_{m'} \;=\; \tilde{k}_m\, \tilde{k}_{m'} \qquad , \tag{3.22}$$

and gives a meaningful square of (3.10) (we will see in a moment that $\tilde{k}_m\, \tilde{k}_{m'}$ vanishes for $m \neq m'$). If our construction ensures that $\tilde{p}_m$ is a positive operator, (3.22) is even equivalent to (3.10).

We calculate the product $\tilde{k}_m\, \tilde{k}_{m'}$ explicitly: The definitions (3.6),(3.7) and (3.11),(2.6) yield in analogy to (3.19) the formulas

$$p_m\, s_{m'} \;=\; s_{m'}\, p_m \;=\; \mathrm{PP}\left(\frac{1}{m-m'}\right) p_m \tag{3.23}$$

$$k_m\, s_{m'} \;=\; s_{m'}\, k_m \;=\; \mathrm{PP}\left(\frac{1}{m-m'}\right) k_m \tag{3.24}$$

$$s_m\, s_{m'} \;=\; \mathrm{PP}\left(\frac{1}{m-m'}\right)(s_m - s_{m'}) \qquad , \tag{3.25}$$

where $\mathrm{PP}(x^{-1}) = \frac{1}{2}\lim_{0<\varepsilon\to0}[(x+i\varepsilon)^{-1} + (x-i\varepsilon)^{-1}]$ denotes the principal value. As a consequence, the operator products with factors $s_m, s_{m'}$ are telescopic, i.e.

$$\sum_{p=0}^{n} k_m\, (\mathcal{B}\, s_m)^p\, (s_{m'}\, \mathcal{B})^{n-p}\, k_{m'} \;=\; 0 \qquad \text{for } n \geq 1. \tag{3.26}$$

This allows us to easily carry out the product $b_m^>\, b_m^<$ in the expression

$$k_m\, b_m^>\, b_{m'}^<\, k_{m'} \;=\; \delta(m-m')\, p_m \qquad . \tag{3.27}$$

With this formula, we can calculate the square of (3.12) to

$$\tilde{k}_m\, \tilde{k}_{m'} \;=\; \delta(m-m') \sum_{\beta_1,\beta_2=0}^{\infty} (-i\pi)^{2\beta_1+2\beta_2}\, b_m^<\, (k_m\, b_m)^{2\beta_1}\, p_m\, (b_m\, k_m)^{2\beta_2}\, b_m^> . \tag{3.28}$$

We could continue the proof by verifying explicitly that the product $\tilde{p}_m\, \tilde{p}_{m'}$ with $\tilde{p}_m$ according to (3.13) coincides with (3.28). This is a straight-forward computation, but it is rather lengthy and not very instructive. We prefer to describe how the operator products (3.16) and the coefficients (3.15) can be derived. In order to keep the proof better readable, we make some simplifications: Since the factors $b_m^<, b_m^>$ cancel similar to (3.27) in telescopic sums, we can omit them in all formulas without changing the multiplication rules for the operator products. Then all operator products have k_m or

p_m as their first and last factor, and we can multiply them with the rules (3.19),(3.20), and (3.21). Since all these rules give a factor $\delta(m - m')$, we will in any case get the prefactor $\delta(m - m')$ in (3.28). Therefore we can just forget about all factors $\delta(m - m')$ and consider all expressions at the same value of m. Furthermore, we will omit the subscript '$_m$' and write the intermediate factors b as a dot '.'. After these simplifications, we end up with formal products of the form

$$F_1 \cdot F_2 \cdot F_3 \cdot \cdots \cdot F_n \qquad \text{with} \qquad F_j = k \text{ or } F_j = p \qquad (3.29)$$

and have the multiplication rules

$$p^2 = k^2 = 1 \quad , \qquad p\,k = k\,p = k \qquad . \qquad (3.30)$$

We must find a positive operator $\tilde{p}$ being a formal sum over operator products (3.29) such that

$$\tilde{p}^2 = \sum_{\beta_1,\beta_2=0}^{\infty} (-i\pi)^{2\beta_1+2\beta_2} (k\,.)^{2\beta_1}\, p\,(.\,k)^{2\beta_2} \qquad . \qquad (3.31)$$

In this way, we have reduced our problem to the combinatorics of the operator products. As soon as we have found a solution $\tilde{p}$ of (3.31), the expression for $\tilde{p}_m$ is obtained by adding the subscripts '$_m$' and by inserting the factors $b_m^<$, b_m, $b_m^>$. Relation (3.22) follows as an immediate consequence of (3.31).

The basic step for the calculation of $\tilde{p}$ is to rewrite (3.31) in the form

$$\tilde{p}^2 = p+A \qquad \text{with} \qquad A = \sum_{(\beta_1,\beta_2)\neq(0,0)} (-i\pi)^{2\beta_1+2\beta_2} (k.)^{2\beta_1} p(.k)^{2\beta_2} \quad .$$
$$(3.32)$$

The operator p is idempotent and acts as the identity on A, $Ap = pA = A$. Therefore we can take the square root of $p + A$ with a formal Taylor expansion,

$$\tilde{p} = \sqrt{p + A} = p + \sum_{n=1}^{\infty}(-1)^{n+1}\,\frac{(2n - 3)!!}{n!\,2^n}\,A^n \qquad , \qquad (3.33)$$

which uniquely defines $\tilde{p}$ as a positive operator.

It remains to calculate A^n. If we take the nth power of the sum in (3.32) and expand, we end up with one sum over more complicated operator products. We first consider how these operator products look like: The operator products in (3.32) all contain an even number of factors k and exactly one factor p. The factor p can be the 1st, 3rd,... factor of the product. Each combination of this type occurs in A exactly once. If we multiply n such terms, the resulting operator product consists of a total odd number of factors p, k. It may contain several factors p, which all occur at odd positions in the product. Furthermore, the total number of factors p

is odd, as one sees inductively. We conclude that A^n consists of a sum of operator products of the form

$$(k \, . \, k \, .)^{q_1} \, p \, . \, k \, . \, (k \, . \, k \, .)^{q_2} \, p \, . \, k \, . \, (k \, . \, k \, .)^{q_3} \cdots (k \, . \, k \, .)^{q_{2\alpha+1}} \, p \, (. \, k \, . \, k)^{q_{2\alpha+2}} \quad (3.34)$$

with $\alpha, q_j \geq 0$. We set $\beta = 2\alpha + \sum_j q_j$. Notice that the number of factors p in (3.34) is $2\alpha + 1$; the total number of factors p, k is $2\beta + 1$. The form of the operator product gives the only restriction $0 \leq 2\alpha \leq \beta$ for the choice of the parameters α, β.

Next we count how often every operator product (3.34) occurs in the sum: The easiest way to realize (3.34) is to form the product of the $\alpha + 1$ factors

$$\left[(k.k.)^{q_1} \, p \, (.k.k)^{q_2+1} \right] \left[(k.k.)^{q_3+1} \, p \, (.k.k)^{q_4+1} \right] \cdots \left[(k.k.)^{q_{2\alpha+1}+1} \, p \, (.k.k)^{q_{2\alpha+2}} \right]$$
$$(3.35)$$

However, this is not the only possibility to factorize (3.34). More precisely, we can apply to each factor in (3.35) the identities

$$
\begin{aligned}
(k \, . \, k \, .)^q \, p \, (. \, k \, . \, k)^r &= [(k \, . \, k \, .)^q \, p] \, [p \, (. \, k \, . \, k)^r] \\
(k \, . \, k \, .)^q \, p \, (. \, k \, . \, k)^r &= [(k \, . \, k \, .)^s \, p] \, [(k \, . \, k \, .)^{q-s} \, p \, (. \, k \, . \, k)^r] \\
(k \, . \, k \, .)^q \, p \, (. \, k \, . \, k)^r &= [(k \, . \, k \, .)^q \, p \, (. \, k \, . \, k)^{r-s}] \, [p \, (. \, k \, . \, k)^s]
\end{aligned} \quad .
$$

By iteratively substituting these identities into (3.35), we can realize every factorization of (3.34). Each substitution step increases the number of factors by one. The steps are independent in the sense that we can fix at the beginning at which positions in (3.35) the product shall be split up, and can then apply the steps in arbitrary order. There are $(\alpha + 1) + (q_1 - 1) + \sum_{j=2}^{2\alpha+1} q_j + (q_{2\alpha+2} - 1) = \beta - (\alpha + 1)$ positions in (3.35) where we could split up the product (in the case $q_1 = 0$ or $q_{2\alpha+2} = 0$, the counting of the positions is slightly different, but yields the same result). Since we want to have n factors at the end, we must choose $n - (\alpha + 1)$ of these positions, which is only possible for $\alpha + 1 \leq n \leq \beta$ and then gives $(\beta - \alpha - 1)! / ((n - \alpha - 1)! \, (\beta - n)!)$ possibilities.

Combining these combinatorial factors with the constraints $0 \leq 2\alpha \leq \beta$, $\alpha + 1 \leq n \leq \beta$ gives for $n \geq 1$

$$
A^n = \sum_{\beta=n}^{\infty} \sum_{\alpha=0}^{\min\left(n-1, \left[\frac{\beta}{2}\right]\right)} \binom{\beta - \alpha - 1}{n - \alpha - 1} \sum_{Q \in \mathcal{P}(\beta+1), \ \#Q = 2\alpha+1}
$$
$$
\times (-i\pi)^{2\beta} \, F(Q, 1) \, . \, k \, . \, F(Q, 2) \, . \, k \, . \, \cdots \, . \, k \, . \, F(Q, \beta+1) \quad (3.36)
$$

with $F(Q, n) = p$ for $n \in Q$ and $F(Q, n) = k$ otherwise. Notice that the last sum in (3.36) runs over all possible configurations of the factors p, k in the operator product (3.34) for fixed α, β. We finally substitute this formula into (3.33) and pull the sums over α, β outside. This gives the desired formula

for $\tilde{p}$.   ∎

In order to illustrate the derived formulas for $\tilde{p}$ and $\tilde{k}$, we give the contribution up to third order in more detail:

$$
\begin{aligned}
\tilde{k}_m \;=\; & k_m \;-\; k_m\,\mathcal{B}\,s_m \;-\; s_m\,\mathcal{B}\,k_m \\
& + s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,k_m \;+\; s_m\,\mathcal{B}\,k_m\,\mathcal{B}\,s_m \;+\; k_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m \;-\; \pi^2\,k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m \\
& - s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,k_m \;-\; s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,k_m\,\mathcal{B}\,s_m \\
& \qquad - s_m\,\mathcal{B}\,k_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m \;-\; k_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m \\
& + \pi^2\,s_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m \;+\; \pi^2\,k_m\,\mathcal{B}\,s_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m \\
& \qquad + \pi^2\,k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,s_m\,\mathcal{B}\,k_m \;+\; \pi^2\,k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,s_m \;+\; \mathcal{O}(\mathcal{B}^4) \\
\tilde{p}_m \;=\; & p_m \;-\; p_m\,\mathcal{B}\,s_m \;-\; s_m\,\mathcal{B}\,p_m \\
& + s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,p_m \;+\; s_m\,\mathcal{B}\,p_m\,\mathcal{B}\,s_m \;+\; p_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m \\
& - \frac{\pi^2}{2}\,p_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m \;-\; \frac{\pi^2}{2}\,k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,p_m \\
& - s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,p_m \;-\; s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,p_m\,\mathcal{B}\,s_m \\
& \qquad - s_m\,\mathcal{B}\,p_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m \;-\; p_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m\,\mathcal{B}\,s_m \\
& + \frac{\pi^2}{2}(s_m\,\mathcal{B}\,p_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m \;+\; p_m\,\mathcal{B}\,s_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m \\
& \qquad + p_m\,\mathcal{B}\,k_m\,\mathcal{B}\,s_m\,\mathcal{B}\,k_m \;+\; p_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,s_m) \\
& + \frac{\pi^2}{2}(s_m\,\mathcal{B}\,k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,p_m \;+\; k_m\,\mathcal{B}\,s_m\,\mathcal{B}\,k_m\,\mathcal{B}\,p_m \\
& \qquad + k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,s_m\,\mathcal{B}\,p_m \;+\; k_m\,\mathcal{B}\,k_m\,\mathcal{B}\,p_m\,\mathcal{B}\,s_m) \;+\; \mathcal{O}(\mathcal{B}^4) \qquad .
\end{aligned}
$$

The theorem gives precise formulas for the perturbation expansion of the Dirac sea. Both the combinatorics of the factors k_m, p_m, s_m and the numerical prefactors are a non-trivial result and, as far as the author knows, cannot be understood intuitively.

We call the perturbation expansion of this theorem the *causal perturbation expansion*. It allows to uniquely define the Dirac sea by

$$
\tilde{P}(x,y) \;=\; \frac{1}{2}\,(\tilde{p}_m - \tilde{k}_m)(x,y) \qquad .
$$

4 Generalization to Systems of Dirac Seas

In the previous section, we defined the Dirac sea for a system of interacting fermions of mass m. A realistic model, however, is composed of several types of fermionic particles with masses $m_1, \ldots, m_f$. Furthermore, the fermions of zero mass may (like the neutrinos in the standard model) occur only as left or right handed particles. The perturbation $\mathcal{B}$ will in general mix up the eigenstates to different masses and will in this way describe an interaction of

all the fermions. We will now extend the previous construction to this more general setting.

First we must generalize (1.3) and define a distribution $P(x, y)$ which describes the system in the vacuum: In order to distinguish the chirality of the zero-mass fermions, we introduce (4×4)-matrices $X_1, \ldots, X_f$. For the zero-mass fermions $m_j = 0$, they can be either $X_j = \mathbb{1}$, $X_j = \chi_L$, or $X_j = \chi_R$, where $\chi_{L/R} = \frac{1}{2}(1 \mp \gamma^5)$ are the chiral projectors. For $m_j \neq 0$, they must coincide with the identity $X_j = \mathbb{1}$. The Dirac seas of the individual types of fermions are then described by $X_j \frac{1}{2}(p_{m_j} - k_{m_j})(x, y)$. The remaining question is how to build up $P(x, y)$ from the individual Dirac seas. In view of the configuration and the interactions of the fermions in the standard model, one might want to use combinations of sums and direct sums

$$P(x, y) = \bigoplus_l \sum_\alpha X_{l\alpha} \frac{1}{2} (p_{m_{l\alpha}} - k_{m_{l\alpha}})(x, y) \tag{4.1}$$

(e.g. with $l = 1, \ldots, 8 = 2(3 + 1)$ running over the color, lepton, and isospin index, and with the index $\alpha = 1, \ldots, 3$ to distinguish the three fermion families. It seems reasonable to use the ordinary sum over α because the families show the same interactions). From the mathematical point of view, however, it is easier to use only direct sums

$$P(x, y) = \bigoplus_{l=1}^f X_l \frac{1}{2}(p_{m_l} - k_{m_l})(x, y) \qquad . \tag{4.2}$$

This is no loss of generality, because the ansatz (4.1) can be obtained from (4.2) by taking a suitable partial trace over the l-index (in our example, by choosing $f = 24 = 3 \cdot 8$ and forming the trace over the three families). For the perturbation expansion, we can also restrict ourselves to the ansatz (4.2), because the perturbation expansion for (4.1) is obtained by taking the partial trace of $\tilde{P}(x, y)$ (see [4] for a more detailed discussion of this method). Therefore we must in the following only consider a $P(x, y)$ of the form (4.2); it is called the *fermionic projector of the vacuum*.

It is convenient to use a matrix notation in the direct sum: We set

$$p(x, y) = \oplus_{l=1}^f p_{m_l}(x, y) \quad , \qquad k(x, y) = \oplus_{l=1}^f k_{m_l}(x, y)$$

and define the matrices

$$X = \bigoplus_{l=1}^f X_l \quad , \qquad Y = \frac{1}{m} \bigoplus_{l=1}^f m_l \qquad ,$$

which are called *chiral asymmetry matrix* and *mass matrix*, respectively (m is an arbitrary mass parameter; e.g. one can choose $m = \max_j m_j$). Then we can write the fermionic projector as

$$P(x, y) = X \frac{1}{2} (p(x, y) - k(x, y)) \qquad . \tag{4.3}$$

Since $m_j = 0$ for $X \neq \mathbb{1}$ and because $p_{m=0}$, $k_{m=0}$ anti-commute with γ^5, we have alternatively

$$P(x,y) \;=\; \frac{1}{2}\left(p(x,y) - k(x,y)\right) X^* \qquad , \qquad (4.4)$$

where $X^* = \gamma^0 X^\dagger \gamma^0$ is the adjoint with respect to the scalar product (1.4). The fermionic projector is a solution of the free Dirac equation

$$(i\partial\!\!\!/_x - mY)\, P(x,y) \;=\; 0 \qquad .$$

In order to describe the interacting system, we again insert a differential operator $\mathcal{B}$ into the Dirac equation. Thus the fermionic projector $\tilde{P}(x,y)$ is supposed to be a solution of the Dirac equation

$$(i\partial\!\!\!/_x + \mathcal{B} - mY)\, \tilde{P}(x,y) \;=\; 0 \qquad . \qquad (4.5)$$

$\mathcal{B}$ may be non-diagonal in the "Dirac sea index" l; we assume it to be Hermitian with respect to the scalar product

$$<\Psi \,|\, \Phi> \;=\; \sum_{l=1}^{f} \int \overline{\Psi_l(x)}\, \Phi_l(x)\, d^4x \qquad .$$

The perturbation expansion for k and p can be carried out exactly as in the previous section: We define the advanced and retarded Green's functions by

$$s^\vee(x,y) \;=\; \oplus_{j=1}^{f}\, s_{m_j}^\vee(x,y) \quad , \qquad s^\wedge(x,y) \;=\; \oplus_{j=1}^{f}\, s_{m_j}^\wedge(x,y) \qquad .$$

Their perturbation expansion is, in analogy to (3.2),(3.3), uniquely given by

$$\tilde{s}^\vee \;=\; \sum_{k=0}^{\infty}(-s^\vee \mathcal{B})^k\, s^\vee \quad , \qquad \tilde{s}^\wedge \;=\; \sum_{k=0}^{\infty}(-s^\wedge \mathcal{B})^k\, s^\wedge \qquad . \qquad (4.6)$$

Theorem 4.1 *The perturbation expansion for p and k is uniquely determined by the conditions*

$$\tilde{k} \;=\; \frac{1}{2\pi i}\left(\tilde{s}^\vee - \tilde{s}^\wedge\right) \quad , \qquad \tilde{p} \stackrel{formally}{=} \sqrt{\tilde{k}^2} \qquad . \qquad (4.7)$$

We have the explicit formulas

$$\tilde{k} \;=\; \sum_{\beta=0}^{\infty}(-i\pi)^{2\beta}\, b^<\, k\, (b\,k)^{2\beta}\, b^> \quad , \qquad \tilde{p} \;=\; \sum_{\beta=0}^{\infty}\sum_{\alpha=0}^{\left[\frac{\beta}{2}\right]} c(\alpha,\beta)\, G(\alpha,\beta)$$

with

$$c(0,0) \;=\; 1 \qquad,$$

$$c(\alpha,\beta) \;=\; \sum_{n=\alpha+1}^{\beta} (-1)^{n+1} \frac{(2n-3)!!}{n!\,2^n} \begin{pmatrix} \beta-\alpha-1 \\ n-\alpha-1 \end{pmatrix} \qquad \textit{for } \beta \geq 1 \textit{ and}$$

$$G(f,g) \;=\; \sum_{Q\in\mathcal{P}(\beta+1),\ \#Q=2\alpha+1} (-i\pi)^{2\beta}\, b^{<}\, F(Q,1)\, bkb\, F(Q,2)\, bkb \cdots$$

$$\times bkb\, F(Q,\beta+1)\, b^{>} \quad,$$

where $\mathcal{P}(n)$ *is the set of subsets of* $\{1,\dots,n\}$ *and where we used the notation*

$$s \;=\; \frac{1}{2}\left(s^{\vee}+s^{\wedge}\right) \;,\qquad F(Q,n) \;=\; \begin{cases} p & \textit{if } n\in Q \\ k & \textit{if } n\notin Q \end{cases}$$

$$b^{<} \;=\; \sum_{k=0}^{\infty}(-s\,\mathcal{B})^k \;,\qquad b \;=\; \sum_{k=0}^{\infty}(-\mathcal{B}\,s)^k\,\mathcal{B} \;,\qquad b^{>} \;=\; \sum_{k=0}^{\infty}(-\mathcal{B}\,s)^k.$$

Proof: Follows exactly as Theorem 3.2. ∎

After this straightforward generalization, we come to the more interesting question of how $\tilde{P}$ can be defined. Our first idea is to set in generalization of (4.3)

$$\tilde{P}(x,y) \;=\; X\,\frac{1}{2}\,(\tilde{p}-\tilde{k})(x,y) \qquad . \tag{4.8}$$

This is not convincing, however, because we could just as well have defined $\tilde{P}(x,y)$ in analogy to (4.4) by $\tilde{P} = \frac{1}{2}(\tilde{p}-\tilde{k})\,X^{*}$, which does not coincide with (4.8) as soon as X, X^{*} do not commute with $\mathcal{B}$. It turns out that this arbitrariness in defining the Dirac sea reflects a basic problem of the causal perturbation expansion for systems with chiral asymmetry. In order to describe the problem in more detail, we consider the perturbation calculation for k to first order: According to (4.6),(4.7), we have

$$\begin{aligned} \tilde{k} \;&=\; k \;-\; \frac{1}{2\pi i}\,(s^{\vee}\,\mathcal{B}\,s^{\vee} - s^{\wedge}\,\mathcal{B}\,s^{\wedge}) \;+\; \mathcal{O}(\mathcal{B}^2) \\ &=\; k \;-\; s\,\mathcal{B}\,k \;-\; k\,\mathcal{B}\,s \;+\; \mathcal{O}(\mathcal{B}^2) \qquad . \end{aligned} \tag{4.9}$$

This expansion is causal in the sense that $\tilde{k}(x,y)$ only depends on $\mathcal{B}$ in the "diamond" $(L_x^{\vee}\cap L_y^{\wedge})\cup(L_y^{\vee}\cap L_x^{\wedge})$, as is obvious in (4.9). It is not clear, however, how to insert the chiral asymmetry matrix into this formula. It seems most natural to replace all factors k by Xk,

$$(\tilde{Xk}) \;=\; Xk \;-\; s\,\mathcal{B}\,Xk \;-\; Xk\,\mathcal{B}\,s \;+\; \mathcal{O}(\mathcal{B}^2) \qquad . \tag{4.10}$$

This formula really gives a possible perturbation expansion for the system of Dirac seas. Unfortunately, it cannot be expressed similar to (4.9) with the

advanced and retarded Green's functions, which means that the causality of the expansion is in general lost. In order to avoid this problem, one might want to insert X at every factor s, k,

$$
\begin{aligned}
(\tilde{X}k) &= Xk - Xs\,\mathcal{B}\,Xk - Xk\,\mathcal{B}\,Xs + \mathcal{O}(\mathcal{B}^2) \\
&= Xk - \frac{1}{2\pi i}\left(Xs^\vee\,\mathcal{B}\,Xs^\vee - Xs^\wedge\,\mathcal{B}\,Xs^\wedge\right) + \mathcal{O}(\mathcal{B}^2) \qquad (4.11)
\end{aligned}
$$

Similar to (4.9), this expansion is causal. In general, however, it does not give a solution of the Dirac equation $(i\partial\!\!\!/ + \mathcal{B} - m)\,\tilde{k} = 0$, which does not make sense.

The only way out of this dilemma is to impose that the perturbation expansions (4.10) and (4.11) must coincide. This yields a condition on the perturbation operator $\mathcal{B}$, which can be characterized as follows: We demand that

$$
Xs^\vee\,\mathcal{B}\,Xs^\vee = s^\vee\,\mathcal{B}\,Xs^\vee = Xs^\vee\,\mathcal{B}\,s^\vee \qquad . \qquad (4.12)
$$

Since the operator $s^\vee_{m=0}$ anti-commutes with γ^5, we have $Xs^\vee = s^\vee X^*$. Substituting into the second equation of (4.12) yields the condition $X^*\mathcal{B} = \mathcal{B}\,X$. Since X is idempotent, this condition automatically implies the first equation of (4.12). We formulate the derived condition for the whole Dirac operator $i\partial\!\!\!/ + \mathcal{B} - mY$ and thus combine it with the fact that chiral fermions are massless (i.e. $X^*Y = YX = Y$) and that X is composed of chiral projectors (which implies that $X^*\partial\!\!\!/ = \partial\!\!\!/X$).

Def. 4.2 *The Dirac operator is called* **causality compatible** *with X if*

$$
X^*\left(i\partial\!\!\!/ + \mathcal{B} - mY\right) = \left(i\partial\!\!\!/ + \mathcal{B} - mY\right)X \qquad . \qquad (4.13)
$$

In the perturbation expansion to higher order, the condition (4.13) allows to commute X through all operator products. Using idempotence $X^2 = X$, we can moreover add factors X to the product, especially

$$
X\,C_1\,\mathcal{B}\,C_1\,\mathcal{B}\,\cdots\,\mathcal{B}\,C_n = XC_1\,\mathcal{B}\,XC_1\,\mathcal{B}\,\cdots\,\mathcal{B}\,XC_n
$$

with

$$
C_j = p,\; C_j = k \text{ or } C_j = s \quad .
$$

This ensures the causality of the perturbation expansion. For a Dirac operator which is causality compatible with X, the *fermionic projector in the external field* is uniquely defined by (4.8).

5   Discussion, Outlook

In this paper, we gave the formal definition of the Dirac sea in the presence of external fields. The method differs considerably from earlier attempts

to solve the external field problem (see e.g. [2] and the references therein). Namely, in these previous approaches, the Dirac sea was always constructed as the "negative frequency solutions" of the Dirac equation. The basic problem of this concept is that the notions of "positive" and "negative" frequency do not make sense in the case with general interaction. Therefore the construction was always limited to potentials which are either static or have an only adiabatic time dependence. As shown in this paper, the notion of "negative frequency states" is not essential for the description of the Dirac sea. For a general definition of the Dirac sea, it must be given up and must be replaced by a suitable causality condition. In this way, it becomes possible to define the Dirac sea in the presence of potentials with arbitrary time dependence. Although the details of the perturbation expansion are a bit complicated, the basic concept is very simple. The construction is explicitly covariant. It puts the usual "hole"-interpretation of the Dirac equation on a satisfying theoretical basis.

In order to clarify the connection to the usual definition of the Dirac sea, we describe how our definition simplifies in the limit of static potentials: If considered as multiplication operators, static potentials map functions of positive (negative) frequency into functions of positive (negative) frequency. Since p, k, and s are diagonal in momentum space, they clearly also preserve the sign of the frequency. Thus we have

$$[P^{\pm}, p] \;=\; [P^{\pm}, k] \;=\; [P^{\pm}, s] \;=\; [P^{\pm}, \mathcal{B}] \;=\; 0 \qquad , \qquad (5.1)$$

where $P^{\pm}$ denote the projectors on the states of positive and negative frequency, respectively. The operators p and k only differ by a relative minus sign for the states of positive and negative frequency,

$$P^{\pm} \, p \;=\; \pm \, P^{\pm} \, k \qquad .$$

Using this relation together with (5.1), we can replace pairs of factors p by pairs of factors k, e.g.

$$
\begin{aligned}
\cdots p\,\mathcal{B}\cdots p\,\mathcal{B}\cdots \;&=\; \cdots p\,\mathcal{B}\cdots p\,\mathcal{B}\cdots (P^+ + P^-) \\
&=\; P^+ (\cdots k\,\mathcal{B}\cdots k\,\mathcal{B}\cdots) \;+\; P^- (\cdots (-k)\,\mathcal{B}\cdots (-k)\,\mathcal{B}\cdots) \\
&=\; \cdots k\,\mathcal{B}\cdots k\,\mathcal{B}\cdots \qquad ,
\end{aligned}
$$

where the dots '$\cdots$' denote any combination of the operators s, k, p, and $\mathcal{B}$. This allows us to simplify the formula for $\tilde{p}$ by only using exactly one factor p in every operator product. After going through the details of the combinatorics, one obtains the formula

$$\tilde{p} \;=\; \sum_{b=0}^{\infty} (-i\pi)^{2b}\, b^{<}\, p\, (b\,k)^{2b}\, b^{>} \qquad .$$

Thus the Dirac sea (4.8) can be written as

$$\tilde{P}(x,y) \;=\; \sum_{b=0}^{\infty} (-i\pi)^{2b}\, b^{<} \left[\frac{1}{2}\, X\, (p-k)\right] (b\,k)^{2b}\, b^{>} \qquad .$$

This equation shows that $\tilde{P}(x, y)$ is composed of the negative-frequency eigenstates of the Dirac operator (notice that the expression in the brackets $[\cdots]$ is the fermionic projector of the vacuum and that all other factors preserve the sign of the frequency). Thus, for static potentials, our definition is equivalent to the usual concept of "negative frequency states." On the other hand, this consideration illustrates in which way our definition goes beyond the usual picture.

In order to get a better understanding of the time-dependent situation, we next consider a scattering process. For simplicity, we use the elementary framework of [1], but our consideration also applies to the operator algebra and Fock space formalism as e.g. described in [5]. We first recall how a scattering process is commonly described in the classical Dirac theory. We assume the scattering to take place in finite time $t_0 < t < t_1$. This means that the external perturbation $\mathcal{B}$ in (1.1) vanishes outside this time interval,

$$\mathcal{B}(t, \vec{x}) = 0 \qquad \text{for} \qquad t \notin [t_0, t_1] \quad . \tag{5.2}$$

We consider a solution $\tilde{\Psi}$ of the Dirac equation with interaction (1.1). According to (5.2), $\tilde{\Psi}(t, \vec{x})$ is, for $t < t_0$, a solution of the free Dirac equation. We uniquely extend this free solution to the whole Minkowski space and denote it by $\tilde{\Psi}_{\text{in}}$, i.e.

$$(i\partial\!\!\!/ - m)\, \tilde{\Psi}_{\text{in}} = 0 \qquad \text{with} \qquad \tilde{\Psi}_{\text{in}}(t, \vec{x}) = \tilde{\Psi}(t, \vec{x}) \quad \text{for } t < t_0.$$

Similarly, $\tilde{\Psi}(t, \vec{x})$ is also for $t > t_1$ a solution of the free Dirac equation; we denote its extension by $\tilde{\Psi}_{\text{out}}$,

$$(i\partial\!\!\!/ - m)\, \tilde{\Psi}_{\text{out}} = 0 \qquad \text{with} \qquad \tilde{\Psi}_{\text{out}}(t, \vec{x}) = \tilde{\Psi}(t, \vec{x}) \quad \text{for } t > t_1.$$

The wave functions $\tilde{\Psi}_{\text{in}}$ and $\tilde{\Psi}_{\text{out}}$ are called the incoming and outgoing scattering states. The S-matrix S maps the incoming scattering states into the corresponding outgoing states, i.e.

$$\tilde{\Psi}_{\text{out}} = S\, \tilde{\Psi}_{\text{in}} \qquad \text{for every } \tilde{\Psi} \text{ with } (i\partial\!\!\!/ + \mathcal{B} - m)\, \tilde{\Psi} = 0 \qquad \quad .$$

As a consequence of the Dirac current conservation, S is a unitary operator (with respect to the scalar product (1.2)). Using the scattering states, one can build up asymptotic Dirac seas for $t < t_0$ and $t > t_1$. Namely, for an observer in the past $t < t_0$, the bosonic potentials are zero. Thus it is natural for him to describe the vacuum with the free Dirac sea (1.3). If this Dirac sea is extended to the whole Minkowski space with external field, one gets the object

$$\tilde{P}^{\wedge}(x, y) = \sum_{a=1,2} \int_{\mathbb{R}^3} \tilde{\Psi}^{\wedge}_{\vec{k}a}(x)\, \overline{\tilde{\Psi}^{\wedge}_{\vec{k}a}(y)}\, d\vec{k} \qquad \qquad ,$$

where the wave functions $\tilde{\Psi}^{\wedge}_{\vec{k}a}$ are the solutions of the perturbed Dirac equation whose incoming scattering states are the plane wave solutions $\Psi_{\vec{k}a}$,

$$(i\partial\!\!\!/ + \mathcal{B} - m)\, \tilde{\Psi}^{\wedge}_{\vec{k}a} = 0 \qquad \text{with} \qquad (\tilde{\Psi}^{\wedge}_{\vec{k}a})_{\text{in}} = \Psi_{\vec{k}a} \qquad \qquad .$$

Accordingly, an observer in the future $t > t_0$ describes the vacuum with the Dirac sea

$$\tilde{P}^\vee(x,y) \;=\; \sum_{a=1,2} \int_{\mathrm{I\!R}^3} \tilde{\Psi}^\vee_{\vec{k}a}(x)\, \overline{\tilde{\Psi}^\vee_{\vec{k}a}(y)}\, d\vec{k} \qquad ,$$

where

$$(i\partial\!\!\!/ + \mathcal{B} - m)\, \tilde{\Psi}^\vee_{\vec{k}a} \;=\; 0 \qquad \text{with} \qquad (\tilde{\Psi}^\vee_{\vec{k}a})_{\text{out}} \;=\; \Psi_{\vec{k}a} \qquad .$$

The states $\tilde{\Psi}^\vee_{\vec{k}a}$ and $\tilde{\Psi}^\wedge_{\vec{k}a}$ have a more explicit form in terms of the perturbation series

$$\tilde{\Psi}^\wedge_{\vec{k}a} \;=\; \sum_{n=0}^{\infty}(-s^\wedge\,\mathcal{B})^n\,\Psi_{\vec{k}a} \qquad \text{and} \qquad \tilde{\Psi}^\vee_{\vec{k}a} \;=\; \sum_{n=0}^{\infty}(-s^\vee\,\mathcal{B})^n\,\Psi_{\vec{k}a} \qquad ,$$

as is immediately verified with (5.2) using that the support of the advanced and retarded Green's functions is the future and past light cone, respectively. The asymptotics of the Dirac seas is completely described by the S-matrix; namely

$$\tilde{P}^\wedge_{\text{in}} \;=\; P \;=\; \tilde{P}^\vee_{\text{out}} \tag{5.3}$$
$$\tilde{P}^\wedge_{\text{out}} \;=\; S\,\tilde{P}^\vee_{\text{out}}\,S^{-1} \tag{5.4}$$
$$\tilde{P}^\vee_{\text{in}} \;=\; S^{-1}\,\tilde{P}^\wedge_{\text{in}}\,. \tag{5.5}$$

The physical scattering process is conveniently described with the two Dirac seas of the observers in the past and in the future: If the physical system is described by $\tilde{P}^\wedge$, for example, the observer in the past is in the vacuum. According to (5), $\tilde{P}^\wedge$ does in general not coincide with the Dirac sea $\tilde{P}^\vee$. This means that for the observer in the future, both positive frequency states are occupied and negative frequency states are unoccupied, so that for him the system contains both particles and anti-particles. This explains the physical effect of pair creation. Other scattering processes are described similarly.

The causal perturbation expansion yields a unique object $\tilde{P}$ describing the Dirac sea in the scattering process. $\tilde{P}$ coincides neither with $\tilde{P}^\vee$ nor with $\tilde{P}^\wedge$; since its construction involves both the advanced and retarded Green's functions, it can be considered as being an "interpolation" between $\tilde{P}^\wedge$ and $\tilde{P}^\vee$. At first sight, it might seem strange that the Dirac sea is now in both asymptotic regions $t < t_0$ and $t > t_1$ described by the same object. Namely, it was essential for our discussion of pair creation that the Dirac seas of the past and future observers were different. It might seem that by redefining the Dirac sea, we no longer have pair creation. Clearly, this is not the case; all physical effects occur in the same way regardless if one works with the asymptotic Dirac seas $\tilde{P}^\wedge$, $\tilde{P}^\vee$ or with $\tilde{P}$. This is because the S-matrix, which completely describes the physical scattering process, does not depend on the definition of the Dirac sea. Thus the choice of the definition of the Dirac sea in the asymptotic regions is merely a matter of

convenience. This may require some explanation: Suppose that we describe the Dirac sea with $\tilde{P}$. Then the asymptotic Dirac seas $\tilde{P}_{\text{in}}$ and $\tilde{P}_{\text{out}}$ consist of both positive and negative frequency states. As a consequence, they are not stable; they tend to decay into the Dirac sea P of all negative-energy states (this is clear physically from the fact that P has lower energy than $\tilde{P}_{\text{in}}$ and $\tilde{P}_{\text{out}}$). Taking this into account, one gets a consistent description of the physical observations. A further complication with $\tilde{P}$ is that the current and energy distributions in the asymptotic regions are in general not homogeneous. For these reasons, it is highly inconvenient to describe the scattering process only with $\tilde{P}$; it is much easier to work with $\tilde{P}^\wedge$ and $\tilde{P}^\vee$. But apart from these purely practical considerations, there is no reason against the description of the Dirac sea with $\tilde{P}$. The great advantage of the causal perturbation expansion is that it gives a unique definition of the Dirac sea, even in the region with interaction $t_0 < t < t_1$. The Dirac sea is not defined with reference to an observer, but becomes a global object of space-time.

Our definition of the Dirac sea is the starting point for the more technical analysis in [6], where all operator products are estimated and computed explicitly in an expansion around the light cone. In order to further clarify the definition of the Dirac sea, we now qualitatively anticipate some results of [6].

First of all, we explain what "causality" of the perturbation expansion for the Dirac sea precisely means: The expansion (3.9) for $\tilde{k}_m(x, y)$ is causal in the strict sense that the perturbation operator $\mathcal{B}(z)$ only enters for z in the "diamond" $z \in (L_x^\vee \cap L_y^\wedge) \cup (L_x^\wedge \cap L_y^\vee)$. Since $p_m(x, y)$ does not vanish for space-like $y - x$, its perturbation expansion, and consequently also the expansion of the Dirac sea, cannot be causal in this strict sense. As is shown in [6], the distribution $\tilde{P}(x, y)$ has singularities on the light cone (i.e. for $(y - x)^2 = 0$). It turns out that these singularities can be completely described in terms of $\mathcal{B}(z)$ and its partial derivatives along the convex line $z \in \overline{xy}$. Our perturbation expansion is causal in this weaker sense. It is even uniquely characterized by this "causality" of the singularities on the light cone.

Both the operator products and the perturbation series were only treated as formal expressions throughout this paper. We outline in which sense these expressions make mathematical sense: It is shown in [6] that all operator products are well-defined distributions if reasonable regularity conditions on $\mathcal{B}$ are assumed. The convergence of the perturbation expansion is a more difficult problem. For chiral and scalar/pseudoscalar potentials, convergence is shown in [6] for the formulas of the light-cone expansion by explicit calculation. For a gravitational field, the situation is more complicated, because the contributions to $\tilde{P}(x, y)$ of higher order in $\mathcal{B}$ become more and more

singular on the light cone. With a Taylor expansion of the δ-distribution

$$\delta(x + a) \;=\; \delta(x) \;+\; a\,\delta'(x) \;+\; \frac{a^2}{2}\,\delta''(x) \;+\; \cdots \qquad ,$$

these contributions can be understood as describing a "deformation" of the light cone (corresponding to the diffeomorphism invariance of General Relativity), but the convergence has not yet been established rigorously.

We finally remark that the fermionic projector $\tilde{P}(x, y)$ of section 4 is considered in [4] as the basic physical object. In this context, the above construction gives a unique characterization of $\tilde{P}$ by a perturbation $\mathcal{B}$ of the Dirac operator. This makes it possible to get a connection to the description of the interaction with classical potentials. It turns out that this "classical limit" is completely determined by the singularities of $\tilde{P}(x, y)$ on the light cone. The "causality" of our perturbation expansion is then directly related to the locality and causality of the classical field equations.

References

[1] J. Bjorken, S. Drell, "Relativistic Quantum Mechanics," *Mc Graw-Hill*, 1964

[2] B. Thaller, "The Dirac Equation," Springer Verlag, 1984

[3] F. Finster, "Local $U(2, 2)$ Symmetry in Relativistic Quantum Mechanics," preprint hep-th/9703083, to appear in J. of Math. Phys.

[4] F. Finster, "Derivation of Field Equations from the Principle of the Fermionic Projector," preprint gr-qc/9606040

[5] A. Ashtekar and A. Magnon, "A Geometric Approach to External Potential Problems in Quantum Field Theory," Gen. Rel. and Grav. 12, 205-223, 1980

[6] F. Finster, "Light-Cone Expansion of the Dirac Sea in the Presence of Chiral and Scalar Potentials," preprint hep-th/9809019