Pauli Exclusion Principle, the Dirac Void and the Preponderance of Matter over Antimatter

G. G. Nyambuya¹

¹National University of Science and Technology, Faculty of Applied Sciences, Department of Applied Physics, Bulawayo, Republic of Zimbabwe

(Dated: April 5, 2016)

In the year 1928, the pre-eminent British physicist – Paul Adrien Maurice Dirac, derived his very successful equation now popularly known as the Dirac equation. This unprecedented equation is one of the most beautiful, subtle, noble and esoteric equations in physics. One of its greatest embellishments is embedded in that this equation exhibits a perfect symmetry – which amongst others – requires, that the Universe contain as much matter as antimatter, or that, for every known fundamental particle, there exists a corresponding antiparticle. We show here that the Dirac theory in its bare form – without the need of the Pauli Exclusion Principle; can – via, its internal logic – beautifully explain the stability of the Dirac Void i.e., the empty Dirac Sea. There is no need for one to ‘uglify’ Dirac’s otherwise beautiful, self-contained and consistent theory by indiscriminately stuffing the Dirac vacuum with an infinite amount of invisible negative energy in-order to prevent the positive energy Electron from falling into the negative energy state.

“A great deal more was hidden in the Dirac equation . . . than the author had expected when he wrote it down in 1928. Dirac himself remarked in one of his talks that . . . his equation was more intelligent than its author. It should be added, however, that it was . . . Dirac who found most of the additional insights.”

– Victor Frederick Weisskopf (1908 – 2002)

INTRODUCTION

As with the Klein-Gordon equation [1, 2], the Dirac equation [3, 4] admits of negative energy solutions. Though no-longer considered a problem today – for some time – their interpretation (negative energies) presented a great deal of difficulty as these negative energy solutions were considered a serious shortcoming of the Dirac theory. For example – from Albert Einstein [5]’s celebrated mass-energy equivalence ($E = mc^2$), negative energy implies negative mass ($m = E/c^2$). According to our mundane understanding of Newton’s Second Law of Motion, a negative mass particle would be accelerated in the opposite direction of an externally applied force [6]. This is not only weird, but absurd! Despite this, physicist have not stopped to imagine or consider the possibility of the existence of negative mass and the consequences thereof [6–12].

In classical physics, a negative energy state causes no trouble because no transition between positive and negative energy states occurs. Therefore – in classical physics, if a particle occupies a positive energy state at any given time, it will never appear cascade down to a negative energy state. The negative energy states are then eliminated as a result of initial conditions stipulating that no such state occurred in the past. In a quantum theory, this device is no longer admissible, as spontaneous emission of radiation can occur as long as a state of lower energy is unoccupied and as long as conservation of angular and linear momenta can be fulfilled. These conservation principles can always be fulfilled under appropriate conditions. There is nothing to prevent an Electron from radiating energy in making a transition to lower and lower states.

In 1930, Dirac [13] resolved the difficulties of interpretation of these ‘nagging’ negative energy states by suggesting his so-called ‘hole’ theory which he prophetically formulated as follows:

“Assume that nearly all the negative energy states are occupied, with one Electron in each state in accordance with the Exclusion Principle of Pauli. The Exclusion Principle makes it impossible for positive energy Electrons to make transition to negative energy states unless they are emptied by some means. Such an unoccupied negative energy state will now appear as something with positive energy, since to make it disappear, i.e. to fill it up, we should have to add an Electron with negative energy. We assume that these unoccupied negative-energy states are the Positrons. The ‘hole’ would have a charge opposite of that of the positive energy particle.”

In an unparalleled triumphant moment for the Dirac
theory, the Positron was experimentally discovered in 1932 by American physicist – Carl David Anderson (1905 – 1991). It has the same spin as the Electron but opposite electrical charge. For this monumental achievement Anderson was awarded the 1936 Nobel Prize in Physics:

\[ \ldots for \ his \ discovery \ of \ the \ Positron. \]

After the direct detection of the Positron, Germany’s great physicist – Werner Karl Heisenberg (1901 – 1976), who made major contributions to Quantum Theory spoke of Dirac’s discovery of antimatter as:

\[ "\ldots perhaps the biggest change of all the big changes in physics of our century \ldots because it changed our whole picture of matter \ldots" \]

It was one of the touchstone and most spectacular consequence of Dirac’s theory that the old concept of the elementary particle based on their stable identity collapsed completely. Sadly, the preponderance of matter over antimatter remains an intriguing puzzle in the otherwise marvellous edifice of the Dirac theory.

Of its endowment – i.e., Dirac’s vacuum out of which the Positron’s existence was foretold: Dirac [13] suggested of it that there be in it – a distribution of Electrons of infinite density everywhere in the Universe and that a perfect vacuum should whence-forth be a region where all the states of positive energy are unoccupied and all those of negative energy are occupied – leading to a crowded Dirac vacuum. However, this infinite distribution does not contribute to the electric field of the vacuum, as, off-course, Maxwell’s [14]’s equation in a perfect vacuum, \((\nabla \cdot \mathbf{E} = 0)\), must hold true. Thus, only departures from the distribution in a vacuum will contribute to the electric charged density of the vacuum. There will be a contributions of \(−e\) for each occupied state of positive energy and a contribution \(+e\) for each unoccupied state of negative energy.

In Dirac’s ‘hole’ theory, the Pauli Exclusion Principle is central and pivotal as it operates so as to prevent a positive-energy Electron from ordinarily making a downward transition to a state of negative energy. It will still be possible, however, for such an Electron to drop into an unoccupied state of negative energy. In this case we should have an Electron and a Positron disappearing simultaneously, their energy being emitted in the form of radiation. The converse process would consist in the creation of an Electron and a Positron from electromagnetic radiation.

What we have presented thus far is only bonafide history of the problem of the interpretation of negative energy solutions in the Dirac theory. The issue of the vacuum, that is, its contents and dynamics is no longer a problem as most of the sticking issues have since been resolved in quantum field theory. In this history that we have presented, we wanted to highlight the importance of the Pauli Exclusion Principle in the stability and dynamics of the vacuum. Dirac evoked this principle in-order to safeguard the \(E\)-positive to \(E\)-negative (and vise-versa) transitions from occurring. It is this that we demonstrate herein that the Dirac does not need the Pauli Exclusion Principle as it can do without it from its own internal logic. We believe, this is a significant result insofar as understanding the Dirac equation is concerned.

**DIRAC EQUATION**

For instructive and completeness purposes, we here write down the Dirac equation. The Dirac equation is given by:

\[ [i\hbar \gamma^\mu \partial_\mu - m_0c] |\psi\rangle = 0, \]  

(1)

where \(|\psi\rangle\) is the Dirac four component wavefunction and:

\[ \gamma^0 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \]  

(2)

are the \(4 \times 4\) Dirac gamma matrices where \(I_2\) and \(0\) are the \(2 \times 2\) identity and null matrices respectively. Throughout this reading, the Greek indices will be understood to mean \(\mu, \nu, ... = 0, 1, 2, 3\) and lower case English alphabet indices \(i, j, k, ... = 1, 2, 3\). In the subsequent section (which is expositive in nature and is presented for latter instructive purposes), we will drop the use of the Dirac Bra-Ket notation and thereafter, we will resume the use of this notation.

**DIRAC FREE PARTICLE SOLUTIONS**

The free particle solutions of the Dirac equation are obtained by assuming a wavefunction of the form \((\psi = u e^{ip \cdot x / \hbar})\) where \(u\) is a four component object, i.e.:

\[ u = \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{pmatrix}, \]  

(3)

Substituting this free particle solution \((\psi = u e^{ip \cdot x / \hbar})\) into (1), one is led to the following set of simultaneous equations:

\[ \begin{align*} 
(E - m_0c^2)u_0 - c(p_x + ip_y)u_3 - cp_z u_2 &= 0 \\
(E - m_0c^2)u_1 - c(p_x + ip_y)u_2 + cp_z u_3 &= 0 \\
(E + m_0c^2)u_2 - c(p_x - ip_y)u_1 - cp_z u_0 &= 0 \\
(E + m_0c^2)u_3 - c(p_x + ip_y)u_0 + cp_z u_1 &= 0 
\end{align*} \]  

(4)

From this one obtains the following two solutions:
\[ \psi(1) = \sqrt{\frac{E + m_0c^2}{2E}} \begin{pmatrix} 1 \\ 0 \\ \frac{c(p_e - ip_s)}{E + m_0c^2} \\ \frac{c(p_e + ip_s)}{E + m_0c^2} \end{pmatrix} e^{+ip_u x^u/h} \]

The factor \( \sqrt{(E + m_0c^2)/2E} \) has been inserted as a normalization constant. These solution \( \psi(1) \) is obtained by setting \( (u_0 = 1; u_1 = 0) \) and then solving for \( u_2 \) and \( u_3 \) and the solution \( \psi(2) \) is obtained by setting \( (u_0 = 0; u_1 = 1) \) and then solving for \( u_2 \) and \( u_3 \). These two solutions \( \psi(1) \) and \( \psi(2) \) are all positive energy solutions and \( \psi(1) \) is a spin-up particle while \( \psi(2) \) is a spin down particle.

The second set of solutions is obtained by assuming a wavefunction of the form \( \psi = |u e^{+ip_u x^u/h}/h \). Substituting this free particle solution \( \psi = |u e^{+ip_u x^u/h}/h \) into (1), one is led to the following set of simultaneous equations:

\[
\begin{align*}
(\!E + m_0c^2\!) u_0 - c(p_x - ip_y) u_3 - c p_y u_2 &= 0 \\
(\!E + m_0c^2\!) u_1 - c(p_x + ip_y) u_2 + c p_y u_3 &= 0 \\
(\!E - m_0c^2\!) u_2 - c(p_x - ip_y) u_1 - c p_y u_0 &= 0 \\
(\!E - m_0c^2\!) u_3 - c(p_x + ip_y) u_0 + c p_y u_1 &= 0
\end{align*}
\]

From this – one obtains the following two solutions:

\[ \psi(3) = \sqrt{\frac{E + m_0c^2}{2E}} \begin{pmatrix} \frac{c(p_e - ip_s)}{E + m_0c^2} \\ \frac{c(p_e + ip_s)}{E + m_0c^2} \\ c(p_e - ip_s) \\ 0 \end{pmatrix} e^{-ip_u x^u/h} \] and \[ \psi(4) = \sqrt{\frac{E + m_0c^2}{2E}} \begin{pmatrix} 0 \\ \frac{c(p_e - ip_s)}{E + m_0c^2} \\ \frac{c(p_e + ip_s)}{E + m_0c^2} \\ 0 \end{pmatrix} e^{-ip_u x^u/h} \]

In the next section – we shall for latter instructive purpose – demonstrate that the Dirac equation is indeed symmetric under charge conjugation. Otherwise this next section together with the previous section, these are both not necessary as these sections present bona-fide knowledge that lays in the public domain. We however, have presented these sections so as to make this reading self-contained and complete by itself.

**DIRAC’S ANTIMATTER**

As is well known – if \( |\psi\rangle \) is a Dirac particle, then, \( (|\psi_c\rangle = \gamma^2 |\psi^\ast\rangle) \) is the anti-Dirac particle. Stated otherwise or in a much clearer manner – if \( |\psi\rangle \) is a particle, then, \( (|\psi_c\rangle = \gamma^2 |\psi^\ast\rangle) \) is the corresponding antiparticle. An antiparticle is one that would behave exactly the same as its particle counterpart when placed in an ambient electromagnetic field that is reversed in comparison to the ambient electromagnetic field in which the particle is immersed. The Physical Law(s) governing \( |\psi\rangle \) and \( |\psi_c\rangle \) are the same. We will demonstrate this by showing that the Dirac equation is symmetric under the reversal of an ambient electromagnetic field.

To demonstrate this symmetry, we proceed as usual, that is, we bring the Dirac particle \( |\psi\rangle \) under the influence of an ambient electromagnetic magnetic field \( A_{\mu}^{\epsilon} \).
(which is a real function). Having done this, the normal procedure of incorporating this ambient electromagnetic magnetic field into the Dirac equation is by making the canonical transformation \( (\partial_\mu \mapsto D_\mu = \partial_\mu + i A^\mu_\text{em}) \), hence equation (1) will now be given by:

\[
[ih\gamma^\mu (\partial_\mu + i A^\mu_\text{em}) - m_0c] |\psi\rangle = 0. \tag{8}
\]

Equation (8) represents the Dirac particle \(|\psi\rangle\) in the im-mersement of an ambient electromagnetic magnetic field. If we are to reverse the particle’s electromagnetic field and that of the ambient electromagnetic magnetic field \( (A^\mu_\text{em} \mapsto - A^\mu_\text{em}) \), then, (8) becomes:

\[
[ih\gamma^\mu (\partial_\mu - i A^\mu_\text{em}) - m_0c] |\psi^*\rangle = 0. \tag{9}
\]

If the Dirac equation is symmetric under charge conjugation, then, there must exist some mathematical transformations, which if applied to (9), would ultimately lead us back to an equation that is equivalent to (8).

To that end – starting from (9), in-order to revert back to (8), the first mathematical operation to be applied to (9) is the complex conjugate operation on the entire equation. So doing, we will have:

\[
[-ih\gamma^\mu (\partial_\mu + i A^\mu_\text{em}) - m_0c] |\psi^*\rangle = 0. \tag{10}
\]

Now – using the fact that:

\[
\gamma^2 \gamma^* = -\gamma^* \gamma^2, \tag{11}
\]

it follows that multiplying equation (10) from the left by \( \gamma^2 \), this will lead us to the equation:

\[
[ih\gamma^\mu (\partial_\mu + i A^\mu_\text{em}) - m_0c] |\psi_c\rangle = 0, \tag{12}
\]

where \(|\psi_c\rangle = (\gamma^2 |\psi^*\rangle)\). Equation (8) and (12) are the same Physical Laws each describing the two particles \(|\psi\rangle\) and \(|\psi_c\rangle\) respectively. Hence our assertion – that if \(|\psi\rangle\) is the particle, then, \(|\psi_c\rangle\) is the antiparticle; is proved.

**QUANTUM PROBABILITY**

Now, according to the set rules of quantum probability calculus developed from the works of the great – German born – British physicist, Max Born (1882–1970) [15], the probability that a particle initially prepared in the state \(|\psi\rangle\) will make a transition to another state (defined by \(|\Psi\rangle\)), this is given by the inner-product \( \langle \Psi | \psi \rangle \). For all conditions of existence, this inner product is not only a positive real number \( i.e. \):

\[
\langle \Psi | \psi \rangle \in \mathbb{R}^+, \tag{13}
\]

but lays between the interval “zero and unity” \( i.e. \):

\[
0 \leq \langle \Psi | \psi \rangle \leq 1. \tag{14}
\]

Whenever \( \langle \Psi | \psi \rangle = 0 \), this means that such a transition \( \langle \psi \mapsto \Psi \rangle \) is strictly forbidden and as such, this transition \( \langle \psi \mapsto \Psi \rangle \) can never take place in the physical Universe; and whenever \( \langle \Psi | \psi \rangle = 1 \), this transition will occur with 100% certainty. On the other hand, if we have \( \langle \Psi | \psi \rangle \notin \mathbb{R}^+ \) and or \( \langle \Psi | \psi \rangle > 1 \) or \( \langle \Psi | \psi \rangle < 0 \), clearly, such transitions are not physical hence not feasible.

Written in full, this inner-product \( \langle \Psi | \psi \rangle \) is defined:

\[
\langle \Psi | \psi \rangle = \int_{x_1}^{x_2} \int_{y_1}^{y_2} \int_{z_1}^{z_2} (\Psi^T)^* \psi \, dx \, dy \, dz, \tag{15}
\]

where \( (\Psi^T)^* \) is transpose conjugate of \( \psi \) and \( (\Psi | \psi) \) is the probability of transition of the particle \(|\psi\rangle\) into the state \(|\Psi\rangle\). This transition will take place in the volume space defined by the physical space boundaries: \( (x := [x_2, x_1]); (y := [y_2, y_1]) \) and \( (z := [z_2, z_1]) \). If we where to consider the entire Universe and assume it to be infinite, then: \( (x := [+\infty, -\infty]); (y := [+\infty, -\infty]) \) and \( (z := [+\infty, -\infty]) \). In the next section, we will evaluate this inner-product and see if it meets the first condition given in equation (13). From this evaluation, our major result shall flow thereof.

**DIRAC & THE PAULI EXCLUSION PRINCIPLE**

What is interesting and constitutes the main point (result) of the present paper is that, if \(|\psi(i)\rangle : i = 1, 2, 3, 4; \) is any of free particle solutions of the Dirac equation and \(|\psi_c(i)\rangle = (\gamma^2 |\psi^*(i)\rangle)\) is the corresponding antiparticle, then:

\[
\langle \psi_c(i) | \psi(i) \rangle \equiv 0. \tag{16}
\]

The meaning of (16) is that the particle-antiparticle transitions \(|\psi(i)\rangle \mapsto |\psi_c(i)\rangle\) are strictly forbidden hence the Pauli Exclusion Principle is not needed to prevent \( E \)-positive to \( E \)-negative transitions. So, Dirac did not really have to look outside for a solution to prevent the particle-antiparticle transitions \( (i.e., E \)-positive to \( E \)-negative transitions).

Apart from the particle-antiparticle transitions we have four \( E \)-positive to \( E \)-negative transitions and these are: \(1\) \(|\psi(1)\rangle \mapsto |\psi(3)\rangle\), represented by the probability \( \langle \psi(1) | \psi(3) \rangle \), \(2\) \(|\psi(1)\rangle \mapsto |\psi(4)\rangle\), represented by the probability \( \langle \psi(1) | \psi(4) \rangle \), \(3\) \(|\psi(2)\rangle \mapsto |\psi(3)\rangle\), \(4\) \(|\psi(2)\rangle \mapsto |\psi(4)\rangle\), \(5\) \(|\psi(3)\rangle \mapsto |\psi(4)\rangle\), \(6\) \(|\psi(3)\rangle \mapsto |\psi(1)\rangle\), \(7\) \(|\psi(4)\rangle \mapsto |\psi(3)\rangle\), \(8\) \(|\psi(4)\rangle \mapsto |\psi(1)\rangle\).
$|\psi(1)\rangle$, represented by the probability $\langle \psi(2)|\psi(3)\rangle$ and (4) $(|\psi(2)\rangle \rightarrow |\psi(4)\rangle)$, represented by the probability $\langle \psi(2)|\psi(4)\rangle$. Below, we evaluate these probabilities. For $\langle \psi(1)|\psi(3)\rangle$, we have:

$$\langle \psi(1)|\psi(3)\rangle = \frac{c(p_x - ip_y)}{E} \int_{y_1}^{y_2} \int_{y_1}^{x_2} dx dy dz. \quad (17)$$

Since $(p_x = 0)$, it follows that $\langle \psi(1)|\psi(3)\rangle \equiv 0$. For $\langle \psi(1)|\psi(4)\rangle$ and $\langle \psi(2)|\psi(3)\rangle$, we have:

$$\langle \psi(1)|\psi(4)\rangle = \frac{c(p_x + ip_y)}{E} \int_{y_1}^{y_2} \int_{y_1}^{x_2} dx dy dz, \quad (18)$$

and:

$$\langle \psi(2)|\psi(3)\rangle = \frac{c(p_x + ip_y)}{E} \int_{y_1}^{y_2} \int_{y_1}^{x_2} dx dy dz. \quad (19)$$

Since $(p_x \neq 0)$ and $(p_y \neq 0)$, then, in both cases (18 and 19), we have $\langle \psi(1)|\psi(4)\rangle \in \mathbb{C}$ and $\langle \psi(2)|\psi(3)\rangle \in \mathbb{C}$), the meaning of which is that these transitions are not feasible. For $\langle \psi(2)|\psi(4)\rangle$, we have $\langle \psi(2)|\psi(4)\rangle \equiv 0$, the meaning of which is that this transition is strictly and completely forbidden. In summary, particle-antiparticle transitions ($\psi \rightarrow \psi_\ast$) are strictly and completely forbidden while $E$-positive to $E$-negative transitions are also not allowed. Hence, the Dirac Void does not need the Pauli Exclusion Principle to stabilize it. The Dirac Void is simply an empty Dirac Sea.

**GENERAL DISCUSSION**

While the free particle solutions of the Dirac equation have been known since the Dirac equation was discovered (88 years ago), it strongly appears to us – that – simple as it may, the result obtained in the present reading may well be new. For this where not the case, it certainly would be a popular and **bona-fide** result as this would have lead to the (beautiful) discarding of the Pauli Exclusion Principle in midst of the Dirac theory where this principle has ‘painfully’ been used to try to bring about harmony in the Dirac World. The Dirac theory in its bare form – without the Pauli Exclusion Principle; can beautifully explain the stability of the Dirac Void. There was – and there, is no need for Dirac (or anyone for that matter) to ‘uglify’ his (Dirac’s) otherwise beautiful, self-contained and consistent theory by indiscriminately stuffing the Dirac vacuum with an infinite amount of negative energy in order to prevent the positive energy Electron from falling into the negative energy state.

At the end of the day, the result that we have is: $\langle \psi|\psi\rangle = 0$, which means $E$-positive to $E$-negative transitions are not allowed. Apart from this – this result $\langle \psi|\psi\rangle = 0$, also means that the matter and antimatter states do not overlap, the deeper meaning of which is that both matter and antimatter can not be found in the region of space or Universe. Just as two Fermionic quantum states can not overlap – *i.e.*, can not be found in the same region of space (Universe), matter and antimatter can not harmoniously reside in the same Universe. The Universe can only contain one form of material – *i.e.*, either, it contains matter or antimatter and never both.

**Conclusion**

1. The Pauli Exclusion Principle is not necessary in-order to prevent the $E$-positive states from transmuting to $E$-negative states. There really is no need for an exogenous explanation as the internal logic of the Dirac theory can handle this.

2. The ponderous cosmic mystery of the preponderance of matter over antimatter which is widely believed to be a shortcoming of the Dirac theory is well explained by the Dirac theory well within the internal logic of the theory since the theory boldly predicts a non-overlap between matter and antimatter.

* Electronic address: golden.nyambuya@nust.ac.zw, physicis@gmail.com