

A beautiful sea: P. A. M. Dirac's epistemology and ontology of the vacuum

Aaron Sidney Wright

To cite this article: Aaron Sidney Wright (2016) A beautiful sea: P. A. M. Dirac's epistemology and ontology of the vacuum, *Annals of Science*, 73:3, 225-256, DOI: [10.1080/00033790.2016.1157731](https://doi.org/10.1080/00033790.2016.1157731)

To link to this article: <http://dx.doi.org/10.1080/00033790.2016.1157731>



Published online: 19 Apr 2016.



Submit your article to this journal [↗](#)



Article views: 47



View related articles [↗](#)



View Crossmark data [↗](#)

A beautiful sea: P. A. M. Dirac's epistemology and ontology of the vacuum

Aaron Sidney Wright 

Department of the History of Science, Harvard University, Cambridge, MA, USA

SUMMARY

This paper charts P.A.M. Dirac's development of his theory of the electron, and its radical picture of empty space as an almost-full plenum. Dirac's Quantum Electrodynamics famously accomplished more than the unification of special relativity and quantum mechanics. It also accounted for the 'duplexity phenomena' of spectral line splitting that we now attribute to electron spin. But the extra mathematical terms that allowed for spin were not alone, and this paper charts Dirac's struggle to ignore or account for them as a sea of strange, negative-energy, particles with positive 'holes'. This work was not done in solitude, but rather in exchanges with Dirac's correspondence network. This social context for Dirac's work contests his image as a lone genius, and documents a community wrestling with the ontological consequences of their work. Unification, consistency, causality, and community are common factors in explanations in the history of physics. This paper argues on the basis of materials in Dirac's archive that — in addition — mathematical beauty was an epistemological factor in the development of the electron and hole theory. In fact, if we believe that Dirac's beautiful mathematics captures something of the world, then there is both an epistemology and an ontology of mathematical beauty.

ARTICLE HISTORY

Received 2 May 2014

Accepted 20 February 2016

Contents

1. Introduction	225
2. Dirac's vacuum	227
2.1 Responding to Klein and Gordon	227
2.2 Correspondence networks	234
2.3 The anti-electron	238
2.4 A connected Dirac	244
3. Mathematical beauty	245
3.1 Beauty in action	247
3.2 Beauty in retrospect	251
4. Conclusion	253

1. Introduction

For Paul Adrien Maurice Dirac (1902–1984) the vacuum was an infinite sea of strange particles — particles with negative energy. He came to this conclusion on the basis of his relativistic theory of the electron, starting in 1928. And he was led to his theory by beautiful mathematics. Beautiful

mathematics worked for Dirac, he believed, because nature itself was beautiful. This epistemological guide — look first for beautiful mathematics — shaped his specific ontological commitments.

Dirac was unquestionably one of the twentieth century's greatest theoretical physicists. His fame came early, as he was finishing his doctoral studies at Cambridge. He arrived at Cambridge — after pursuing an engineering degree at Bristol — hoping to study relativity. However he was guided toward Ralph Fowler as an advisor, an expert in quantum theory. And in fact, if relativity can be cast as the study of the invariance of physical quantities under certain transformations¹ one observes a continuity in Dirac's work between these two fields. In Copenhagen in 1926 he established a 'general transformation' theory of quantum mechanics that generalized and clarified the work of Schrödinger and Heisenberg.² This work became his 'darling'.³ In 1927 Dirac viewed the dynamics of quantum systems in terms of mathematical transformations that propagated states through time.⁴ According to Helge Kragh, Dirac had 'unshakable confidence in general quantum mechanics, the guiding principle for all of his research from 1926 onwards'.⁵ Though Dirac usually published his work just after quantum mechanics' European innovators, it was ascribed the highest value by that community.⁶ It is perhaps not surprising that his most famous achievement was the creation of a theory to unify the physics of special relativity and quantum theory.⁷ As Laurie Brown has written, if there was a man who made Quantum Electrodynamics (QED) it was surely Dirac.⁸ This paper will examine the dynamics of Dirac's relativistic theory, from its inception in 1928 to its maturity in the mid-1930s. After the War, Dirac's picture would be supplanted by Feynman, Schwinger, and Tomonaga's renormalizable QED.⁹

We can think of Dirac's theory — and perhaps all physical theories — as consisting of interacting parts. There were epistemic strategies and ontological presuppositions. Epistemic strategies are strategies for knowing the world — what forms of investigation generate products that can be warranted as knowledge about the world? Ontological presuppositions are presuppositions about what the world is made of, about what sorts of entities make up reality. This paper charts Dirac's evolving ontological pictures from within the framework of his relativistic theory. These pictures were inextricably linked to Dirac's conception of the vacuum: rather than being true empty space, it was filled by an infinite sea of negative-energy particles. This paper also engages with Dirac's formal, epistemic strategies. Dirac was motivated by the seeming simplicity and beauty of his mathematical theories. My account is indebted to Helge Kragh's scientific biography of Dirac, which charts much of the same territory.¹⁰ (Some of this territory, particularly in section 2.1, has been treated also by D. Monti.¹¹)

A note on 'beauty': I will not attempt to create an understanding of 'beauty' for Dirac more specific than Dirac's own. That is to say, though it may be possible to infer in certain instances a

¹Lorentz transformations for special relativity, and more general transformations in general relativity.

²Olivier Darrigol, *From c-Numbers to Q-Numbers: The Classical Analogy in the History of Quantum Theory* (Berkeley: University of California Press, 1992).

³Cited in Abraham Pais, 'Playing with equations, the Dirac way', in *Paul Dirac: the man and his work*, ed. by Peter Goddard (Cambridge: Cambridge University Press, 1998), pp. 93–116 (pp. 96–97).

⁴P. A. M. Dirac, 'The Physical Interpretation of the Quantum Dynamics', *Proceedings of the Royal Society of London. Series A*, 113, no. 765 (1927), 621–41. doi:10.1098/rspa.1927.0012

⁵Helge Kragh, 'The Genesis of Dirac's Relativistic Theory of Electrons', *Archive for History of Exact Sciences*, 24, no. 1 (1981), 31–67 (p. 50). doi:10.1007/BF00327714

⁶Jagdish Mehra and Helmut Rechenberg, *The Completion of Quantum Mechanics 1926–1941*, vol. 6, *The Historical Development of Quantum Theory 2* (New York: Springer, 2001); Helge Kragh, *Quantum Generations: A History of Physics in the Twentieth Century* (Princeton, NJ: Princeton University Press, 1999).

⁷Kragh, 'The Genesis' (note 5).

⁸Laurie M. Brown, 'Some QED myths-in-the-making? : Silvan S. Schweber, QED and the Men Who Made It: Dyson, Feynman, Schwinger and Tomonaga (Princeton University Press, 1994)', *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, 27, no. 1 (1996), 81–90. doi:10.1016/1355-2198(95)00023-2

⁹S. S. Schweber, *QED and the men who made it: Dyson, Feynman, Schwinger, and Tomonaga* (Princeton, NJ: Princeton University Press, 1994).

¹⁰Helge Kragh, *Dirac: A Scientific Biography* (Cambridge: Cambridge University Press, 1990).

¹¹Dalida Monti, 'Dirac's Holes Model: From Proton to Positron', *Nuncius*, 10, no. 1 (1995), 99–130.

more precise meaning of beauty than a vaguer, general notion, this will not be attempted here. For example, one may look to the aesthetic qualities of the theories Dirac admired to tease out a precise notion of mathematical beauty. For some authors, this has taken the form of noting Dirac's appreciation for, and use of, coordinate transformations. However, Dirac explicitly wrote in his 1939 James Scott lecture that what he meant by beauty was — like beauty in art — resistant to precise definitions (section 3.1).¹² And, Dirac viewed some approximations as beautiful (section 2.3). I take this to contest any neat identification of 'beauty' with a certain 'mathematical structure'. In this sense I am using 'beauty' as an actor's category, not an analyst's category. I do admit that *if* one constructed a more-precise understanding of beauty for Dirac — across the breadth of his career — one could argue that this specific-beauty was more- or less-influential at different points in Dirac's career. This could possibly be connected to the division we now see in Dirac's work in the mid-1930s of earlier, valuable work, and later less-valued work. However, the present paper is focused closely on Dirac's electron theory. The propriety of using mathematical beauty as a guide in scientific investigation may be debated, but that is not my purpose here. My claim is rather conservative: Dirac claimed in retrospect that mathematical beauty played a role in his early work, I think we should take him at his word.

2. Dirac's vacuum

The origin of Dirac's eponymous equation has been ably treated by Kragh.¹³ The 'Dirac equation' that remains in contemporary physics describes the behaviour of electrons in external electromagnetic fields at relativistic velocities. As Kragh notes, there is not much documentary evidence regarding the genesis of Dirac's equation — at least, not in the form of dated materials in his archive at Florida State University. The current paper is not concerned with the genesis of Dirac's equation, but rather with its implications, and the debates it engendered in the community of physicists at the time. This narrative differs from Kragh's biography¹⁴ in that more emphasis is given to Dirac's correspondence and his changing ontological pictures. New readings of archival evidence of Dirac's early commitment to mathematical beauty will be introduced in the latter part of this paper.

2.1. Responding to Klein and Gordon

'It appears that the simplest Hamiltonian for a point-charge electron satisfying the requirements of both relativity and the general transformation theory leads to an explanation of all duality phenomena without further assumption'.¹⁵ This sentence contains the essence of Dirac's approach in his famous paper on the relativistic electron. Consonant with his aesthetics, he searched for the simplest possible elements with which to build a theory.¹⁶ He did this within a Hamiltonian formulation of mechanics. The model of the electron he used was not a spinning sphere, but rather a mathematical point. And he sought to preserve both the requirements of (special) relativity and those of his general approach to quantum mechanical problems. These last two points deserve some elaboration. On Dirac's understanding of the special theory of relativity, time and spacial coordinates should be considered on the same basis. This meant that, for example, if the components of the three-dimensional momentum vector, \vec{p} , were treated as derivatives with respect to *space* in the quantum theory, the time component should be expressed as a derivative with respect to *time*. Symbolically, if $p_x \rightarrow -i\hbar \partial/\partial x$ then $p_0 \rightarrow i\hbar \partial/\partial t$. This p_0 component was identified with the energy of

¹²Darrigol, *From c-Numbers to Q-Numbers* (note 2), p. 302.

¹³Kragh, 'The Genesis' (note 5).

¹⁴Kragh, *Dirac* (note 10), pp. 87–117.

¹⁵P. A. M. Dirac, 'The Quantum Theory of the Electron', *Proceedings of the Royal Society of London, Series A*, 117, no. 778 (1928), 610–24 (p. 610). doi:10.1098/rspa.1928.0023

¹⁶To say that Dirac appealed to an aesthetics of simplicity here is not to equate simplicity with Dirac's concerns for beauty. They are related concepts, but he appealed to them in different contexts and explicitly rejected reducing beauty to simplicity. I thank an anonymous referee for prompting this clarification.

the system, W , in Dirac's notation. (As Kragh notes, this sort of symmetrizing procedure occurred in Dirac's earlier relativistic works as well.¹⁷ See also the paper by Moyer.¹⁸)

This relativistic requirement, alone, was met by Klein's earlier (1926) equation. However, what is now called the Klein-Gordon (K-G) equation did not fit with the general scheme of Dirac's quantum mechanics.¹⁹ What that meant, for example, was that the K-G equation required a special interpretation of the wave function, different from the interpretation of standard quantum mechanics. The problem was that on the K-G scheme, the square modulus of the wave function, $|\psi|^2$, could have negative values. In standard quantum mechanics, however, Max Born's interpretation held that $|\psi|^2$ was interpreted as the probability for some system to be in a given state.²⁰ Thus, under the standard interpretation, the K-G equation would give negative probabilities, which is nonsense. Dirac sought a scheme where Born's probabilistic interpretation would hold in the relativistic case.

Dirac traced the difficulty with the K-G equation to its non-linearity. The K-G equation has second derivatives with respect to time.²¹ However,

The general interpretation of non-relativity quantum mechanics is based on the transformation theory, and is made possible by the wave equation being of the form

$$(H - W)\psi = 0, \quad (1)$$

i.e. being linear in W or $\partial/\partial t$ [...]. The wave equation of the relativity theory must also be linear in W if the general interpretation is to be possible.²²

Non-linearity seemed endemic to relativistic theories of the electron, however. For no electromagnetic field, the wave equation should be given by

$$(-p_0^2 + \mathbf{p}^2 + m^2 c^2)\psi = 0. \quad (2)$$

Recalling the identification of p_0 with the time derivative, the problem of non-linearity is obvious. Dirac puzzled over how to force this quadratic equation into linear form. In retrospect he identified his methodology as playing around with pretty mathematics.²³ Much as quantum mechanics required the introduction of strange numbers that did not follow the standard commutative law of algebra, to linearize equation (2) Dirac required strange new numbers.²⁴ The conditions these numbers, α , must obey were

$$\alpha_\mu^2 = 1, \quad \alpha_\mu \alpha_\nu - \alpha_\nu \alpha_\mu = 0 (\mu \neq \nu), \quad \mu, \nu = 1, 2, 3, 4. \quad (3)$$

Dirac realized that Pauli's spin matrices behaved similarly, but they came only in sets of three. Dirac needed four numbers, represented as matrices, to multiply the terms in equation (2).²⁵ Pauli's matrices were 2×2 matrices; Dirac realized that analogous matrices could be found if he looked for larger 4×4 matrices. Dirac actually defined his α s in terms of six 4×4 matrices in his

¹⁷P. A. M. Dirac, 'Note on the Doppler Principle and Bohr's Frequency Condition', *Proceedings of the Cambridge Philosophical Society*, 22 (1924), 432; Kragh, 'The Genesis' (note 5).

¹⁸Donald Franklin Moyer, 'Origins of Dirac's Electron, 1925–1928', *American Journal of Physics*, 49, no. 10 (1981), 944–9.

¹⁹Walter Gordon, 'Der Comptoneffekt nach der Schrödingerschen Theorie', *Zeitschrift Für Physik*, 40, no. 1/2 (November 1926), 117–77; Oskar Klein, 'Quantentheorie und fünfdimensionale Relativitätstheorie', *Zeitschrift Für Physik*, 37, no. 12 (1926), 895–906. doi:10.1007/BF01397481

²⁰Max Born, 'On the quantum mechanics of collisions', in *Quantum Theory and Measurement*, ed. and trans. by Wojciech Hubert Zurek and John Archibald Wheeler (Princeton, NJ: Princeton University Press, [1926] 1983), pp. 52–61.

²¹There are other senses of mathematical non-linearity, however this is the sense of linearity that was relevant to Dirac and his contemporaries. See C. G. Darwin, 'The Wave Equations of the Electron', *Proceedings of the Royal Society of London. Series A*, 118, no. 780 (1928), 654–80. doi:10.1098/rspa.1928.0076; G. Breit, 'An Interpretation of Dirac's Theory of the Electron', *PNAS*, 14, no. 7 (1928), 553–9.

²²Dirac, 'The Quantum Theory of the Electron' (note 15), p. 612.

²³Pais, 'Playing with equations' (note 3); Thomas Pashby, 'Dirac's Prediction of the Positron: A Case Study for the Current Realism Debate', *Perspectives on Science*, 20, no. 4 (2012), 440–75.

²⁴That is, new to physics.

²⁵Because the vectors in equation (2) have four components.

paper, σ_r and ρ_r ($r = 1, 2, 3$), that could be formed into vectors. This allowed him to write the famous Dirac equation as

$$[p_0 + \rho_1(\boldsymbol{\sigma}, \mathbf{p}) + \rho_3 mc]\psi = 0, \quad (4)$$

where the brackets $(\ , \)$ represent the inner product between the two vectors.²⁶ With this equation Dirac solved the problem of finding a relativistically invariant wave function for the electron that was linear, and so fit with the standard interpretation of quantum mechanics.

However all was not well with the interpretation of Dirac's new equation. In order for equation (4) to make sense, the wave function could not be simply a single function. Instead it had to have four components: $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$. This was a puzzle in itself. Only two components were needed to account for the 'duplexity phenomena' — i.e. the effects of the spin of the electron such as the anomalous Zeeman effect — one for each spin state.²⁷ What to make of the remaining two components of ψ ? At this stage in the interpretation of the equation, Dirac saw that the components referred to solutions where the electron's charge was opposite, $+e$, however, he took that as a sign that the solutions must simply be 'rejected'.²⁸ Dirac did not rush to invest these mathematical terms with physical meaning.

However, others saw an opening for metaphysics. Oskar Klein recalled that Dirac sent a summary of his electron work to Bohr's institute in Copenhagen around Christmas 1927.²⁹ Already the theory was being connected to new ideas about space and time. Niels Bohr wrote to Dirac on 27 February 1928:

Dear Dirac,

I feel very ashamed first now to write and thank you for your great kindness in sending me an abstract of your last most wonderful paper. I need not say, how deeply interested we all are in the new development which it surely will open. Klein has been especially interested in various points concerned with the general ideas of space and time and is very eager to get opportunity to discuss them with you.³⁰

Here Bohr was probably referring to Klein's ongoing work on a five-dimensional theory that attempted to unify *general* relativity and quantum mechanics.³¹

Dirac's former advisor Ralph Fowler communicated the paper to the Royal Society of London on 2 January 1928, and it was published in the *Proceedings* on 1 February. It made an immediate impression in the centres of European physics. From Hamburg, Yoshio Nishina wrote: 'I must congratulate you on your success in obtaining a complete solution of the electron. We are all looking forward to your paper in the Proc. Roy. Soc'. And Werner Heisenberg wrote from Leipzig on 13 February: 'I admire your last work about the spin in the highest degree'.³²

Dirac visited Heisenberg in Leipzig in June 1928 and gave a lecture on his new theory.³³ He motivated his presentation by noting the requirement that in order to have a sensical electric charge density in a theory, the Hamiltonian must be linear in time derivatives. Perhaps at Heisenberg's prompting (see Dirac to Klein 24 July 1928, below), by summer Dirac realized that his $+e$ solutions could not simply be ignored. However, a solution was not yet in sight: at the end of his talk he briefly

²⁶Dirac, 'The Quantum Theory of the Electron' (note 15), p. 615.

²⁷Paul Forman, 'Alfred Landé and the Anomalous Zeeman Effect, 1919-1921', *Historical Studies in the Physical Sciences*, 2 (1970), 153–261.

²⁸'Since half the solutions must be rejected as referring to the charge $+e$ on the electron, the correct number will be left to account for duplexity phenomena'. Dirac, 'The Quantum Theory of the Electron' (note 15), p. 618.

²⁹Interview of Oskar Klein by J. L. Heilbron and L. Rosenfeld on 1963 February 28, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA, www.aip.org/history-programs/niels-bohr-library/oral-histories/4709-4 [accessed 31 May 2013].

³⁰N. Bohr to P. Dirac, 27 February 1928, Series II, Box 1, Folder 6, Paul A.M. Dirac Collection, Florida State University Libraries, Tallahassee, FL, 32306 USA (Hereafter 'Dirac Papers').

³¹Klein, 'Quantentheorie und fünfdimensionale Relativitätstheorie' (note 19); Koray Karaca, 'Historical and Conceptual Foundations of the Higher Dimensional Unification Program in Physics' (PhD thesis, Indiana University, 2010).

³²W. Heisenberg to P. Dirac, 13 February 1928, Dirac Papers, Series II, Box 1, Folder 6.

³³P. A. M. Dirac, 'Über die Quantentheorie des Elektrons', *Physikalische Zeitschrift*, 29, no. 16 (August 1928), 561–63; P. A. M. Dirac, 'Zur Quantentheorie Des Elektrons', in *Leipziger Vorträge 1928: Quantentheorie Und Chemie*, ed. by H. Falkenhagen, trans. by A. Eucken (Leipzig: S. Hirzel, 1928), pp. 85–94.

turned to ‘the difficulties inherent in the current theory that are still far from being solved’.³⁴ He recognized what we would now call the charge symmetry of his equation. ‘If we write the wave equation $-e$ instead of e , one should as a result of the now very different physical problem, get a much different wave equation, while in fact nothing new results’.³⁵ Perhaps the set of four solutions could be divided into two groups, one group for each charge of the electron? This would only work if there were no transitions between the two groups. Unfortunately, such transitions were allowed by the theory. Separating out the $+e$ solutions would not work. ‘However, the probability of these transitions is extremely small (of the order of magnitude $\left(\frac{v}{c}\right)^4$, c = speed of light, v = velocity of the electron)’. Dirac’s conclusion was that ‘Consequently, the present theory is an approximation’.³⁶ Dirac concluded his paper with the sentiment that a major change in outlook would be necessary to fix this problem, and connected it to the metaphysics of space and time. ‘It seems that this problem can be solved only through a fundamental change in our previous idea, and perhaps is related to the difference between past and future’.³⁷ It seems an open question as to whether Dirac was at this stage considering *physical* interpretations of his $+e$ states. The most conservative reading may be that he was not, yet. Rather he saw his theory as a mathematical approximation (valid for a certain range of electron velocities much smaller than c), and saw a need for reworking the entire physical picture at play. Ontological change was on the table, but its direction was not yet clear.

Dirac would not have seen the approximate nature of his theory as a failure. In a 1967 interview with Thomas S. Kuhn, he reflected on his engineering training at Bristol:

I think that this engineering education has influenced me very much in making me learn to tolerate approximations. My natural feelings were to think that only an exact theory would be worth considering. Now, engineers always have to make approximations. I learned that even a theory based on approximations could be a beautiful theory. I rather got to the idea that everything in nature was only approximate, and that one had to be satisfied with approximations, and that science would develop through getting continually more and more accurate approximations, but would never attain complete exactness. I got that point of view through my engineering training, which I think has influenced me very much. As a result of that I haven’t been much interested in questions of mathematical logic or any attempts to form an absolute measure of accuracy, an absolute standard of reasoning. I feel that these things are just not important, that the study of nature through getting ever, [sic] improving approximations is the profitable line of procedure.³⁸

This point will be developed further below, in considering Dirac’s *in situ*, rather than retrospective, discussions of approximation, and in a discussion of mathematical beauty (Section 3).

As Dirac’s work percolated throughout the international community of physicists, the simple abandonment of the $+e$ solutions to Dirac’s equation was further called into question. In particular, Hermann Weyl published two papers from Princeton, N.J., around the middle of 1929 suggesting that these solutions should be identified with the proton.³⁹ There is no indication in Dirac’s archive that he considered the possible physical interpretation of the $+e$ solutions before Weyl’s work. (Only the published Leipzig presentation gives clues to this thinking in this period.) But the issue itself was on his mind. On 24 July 1928, Dirac wrote to Klein: ‘I have not met with any success in my attempts to solve the $\pm e$ difficulty. Heisenberg (whom I met in Leipzig) thinks the problem will not be solved until one has a theory of the proton and electron together’.⁴⁰

By the winter of 1929 we know he was working on a physical interpretation. Bohr wrote on 24 November that

³⁴Dirac, ‘Über die Quantentheorie des Elektrons’ (note 33), p. 563.

³⁵Ibid., p. 563.

³⁶Ibid., p. 563.

³⁷Ibid., p. 563.

³⁸Interview with P. A. M. Dirac by Thomas S. Kuhn and Eugene Paul Wigner, at Wigner’s home, Princeton, New Jersey April 1, 1962, http://www.aip.org/history/ohilist/4575_1.html [accessed 21 October 2013].

³⁹Hermann Weyl, ‘Elektron Und Gravitation. I’, *Zeitschrift Für Physik*, 56, nos. 5-6 (1929), 330–52. doi:10.1007/BF01339504; Hermann Weyl, ‘Gravitation and the Electron’, *PNAS*, 15, no. 4 (1929), 323–34.

⁴⁰Cited in Pais, ‘Playing with equations’ (note 3), p. 99.

From Gamow I hear that you are now back in England again, and that you have made progress with the mastering of the hitherto unsolved difficulties in your theory of the electron. As we have not yet heard about any details Klein and I should be very thankful if you would be so kind to tell us something of your present views.

Bohr's 'view was that the difficulties in your theory might be said to reveal a contrast between the claims of conservation of energy and momentum on one side and of the conservation of the individual particles on the other side'.⁴¹ Dirac was not in the habit of keeping copies of outgoing correspondence, however his well-known reply of 26 November was kept by Bohr.⁴² Dirac started off the physical content of his letter by engaging with Bohr's opinion of the conservation of energy versus conservation of particle number. Dirac's 'own opinion of this question is that [he] should prefer to keep rigorous conservation of energy at all costs and would rather abandon even the concept of matter consisting of separate atoms and electrons than the conservation of energy'.⁴³

Bohr's view can be elaborated with reference to his recently announced interpretation of quantum mechanics, complementarity.⁴⁴ In an unpublished manuscript for the H. H. Wills Memorial Lecture given at the University of Bristol, 5 October 1931, Bohr wrote that space-time descriptions and energy-momentum descriptions were complementary: they could not coexist.

We have thus either space-time description or description where we can use the laws of conservation of energy and momentum. They are complementary to one another. We cannot use them both at the same time. If we want to use the idea of space-time we must have watches and rods which are outside and independent of the object under consideration, in that sense that we have to disregard the interaction between the object and the measuring rod used. We must refrain from determining the amount of momentum that passes into the instrument in order to be able to apply the space-time point of view.⁴⁵

It seems, perhaps, that Bohr and Dirac were talking past one another. Bohr's comment that 'the difficulties in your theory might be said to reveal a contrast between the claims of conservation of energy and momentum on one side and of the conservation of the individual particles on the other side' was not a statement that conservation of energy and momentum were to be abandoned *in toto*, but rather that Bohr saw a sort of complementarity operating. This would be echoed in later work, particularly at Berkeley.⁴⁶

In Dirac's 1929 letter to Bohr, he then moved to discuss his new insight into his theory of the electron: the idea of an infinite distribution of negative-energy electrons, in which the protons could be seen as 'holes'. These negative energy states were the states of positive electric charge introduced by Dirac's equation (4). They could not be simply identified with protons as Weyl might have had it — and as Dirac would emphasize later — because they had negative kinetic energy. A particle with negative kinetic energy would increase in velocity as it *lost* energy, a bizarre behaviour never observed by experiment.

Dirac began his explanation with the generalization of his equation for the case of an arbitrary field:

There is a simple way of avoiding the difficulty of electrons having negative kinetic energy. Let us suppose the wave equation

$$\left[\frac{W}{c} + \frac{e}{c} A_0 + \rho_1 \left(\underline{\sigma}, \underline{p} + \frac{e}{c} \underline{A} \right) + \rho_3 mc \right] \psi = 0 \quad (5)$$

⁴¹N. Bohr to P. Dirac, 24 November 1929, Dirac Papers, Series II, Box 1, Folder 9.

⁴²Kragh, *Dirac* (note 10), pp. 90–3.

⁴³Dirac papers, Series II, Box 14, Folder 2.

⁴⁴Niels Bohr, 'The Quantum Postulate and the Recent Development of Atomic Theory,' *Nature*, 121 (1928), 580–90.

⁴⁵Niels Bohr, 'Space-Time-Continuity and Atomic Physics', in *Niels Bohr Collected Works*, ed. Jørgen Kalckar, vol. 6 Foundations of Quantum Physics I (Amsterdam: North-Holland, 1985), p. 369.

⁴⁶W. H. Furry and J. R. Oppenheimer, 'On the Theory of the Electron and Positive', *Physical Review*, 45, no. 4 (Feb 1934), 245–62. doi:10.1103/PhysRev.45.245

does accurately describe the motion of a single electron. [...]

Let us now suppose there are so many electrons in the world that all these most stable states [negative-energy electrons with high-velocity] are occupied. The Pauli principle will then compel some electrons to remain in less stable states. For example if all the states of $-ve$ energy are occupied and also a few of $+ve$ energy, these electrons with $+ve$ energy will be unable to make transitions to states of $-ve$ energy and will therefore have to behave quite properly. The distribution of $-ve$ electrons will, of course, be of infinite density, but it will be quite uniform so that it will not produce any electromagnetic field and one would not expect to be able to observe it.

This is the first statement of Dirac's picture of the vacuum of electrons as an infinite distribution of negative energy electrons. It was a calm sea that was 'quite uniform,' and unobservable. Dirac continued to integrate protons into his new picture.

It seems reasonable to assume that not all the states of negative energy are occupied, but that there are a few vacancies of 'holes'. Such a hole which can be described by a wave function, like an X-ray orbit would appear experimentally as a thing with $+ve$ energy, since to make the hole disappear (i.e. to fill it up,) one would have to put $-ve$ energy into it. Further, one can easily see that such a hole would move in an electromagnetic field as though it had a $+ve$ charge. These holes I believe to be the protons. When an electron of $+ve$ energy drops into a hole and fills it up, we have an electron and proton disappearing simultaneously and emitting their energy in the form of radiation.⁴⁷

Dirac expressed this view in a paper published on 1 January 1930.⁴⁸ Though in a later interview Dirac claimed that he first thought of an anti-electron, and only afterwards thought of identifying the holes with protons, in print and in his archives, this comes only later.⁴⁹

Dirac's address to Section A of the 1930 meeting of the British Association for the Advancement of Science, held at Bristol, September 3–10, distilled the core of his approach.⁵⁰ His pencil manuscript is preserved in his archive. He began with an explicitly ontological statement about the nature of matter

Matter is made up of atoms, each consisting of a number of electrons (each having a negative charge $-e$) moving round a central nucleus that has a positive charge $+Ne$

It is likely that the atomic nuclei themselves are not simply particles, but are made up of the nuclei of hydrogen atoms (called protons) and the electrons, out of which all matter is built.⁵¹

Dirac, and perhaps his audience, was not happy with this state of affairs, for 'it would be preferable to have all matter built out of one fundamental kind of particle instead of two, electrons and protons'. This was a further reduction of the commonly-held view that there were only two material particles in the universe, the proton and the electron. The aesthetics of simplicity again governed Dirac's thinking. However, Dirac would go on to elucidate a 'connection between the two kinds of particles, electrons and protons, such that they are not independent and there is ultimately only one kind of fundamental particle in nature'. This connection manifested itself in the form of a 'trouble' with the standard picture of the behaviour of particles in mechanics. This trouble, like the discussion of the K-G equation above, was rooted in the non-linearity of energy equations of matter.

The origin of the trouble is that the energy W of a particle is determined in terms of its momentum p according to relativity theory by the equation

$$\frac{W^2}{c^2} - p^2 - m^2c^2 = 0 \quad (6)$$

which is quadratic in W and allows of negative values for W as well as positive ones. [...]

⁴⁷Dirac papers, Series II, Box 14, Folder 2.

⁴⁸P. A. M. Dirac, 'A Theory of Electrons and Protons', *Proceedings of the Royal Society of London. Series A*, 126, no. 801 (1930), 360–5. doi:10.1098/rspa.1930.0013

⁴⁹Kragh, *Dirac* (note 10), p. 96.

⁵⁰Dirac's address took place Monday the 8th at 11am. P. A. M. Dirac, 'The Proton', *Nature*, 126, no. 3181 (October 1930), 605–6.

⁵¹The pencil manuscript is in Dirac Papers, Series II, Box 26, Folder 8. In the manuscript he also notes Oppenheimer's proposal that each of electrons and protons should have their own infinite distribution of negative-energy states, cancelling the infinite charge density. J. R. Oppenheimer, 'Note on the Theory of the Interaction of Field and Matter', *Physical Review*, 35, no. 5 (Mar 1930), 461–77. doi:10.1103/PhysRev.35.461

Of course a particle with negative energy (kinetic energy is here referred to) has no physical meaning. Such a particle would have less energy the faster it is moving and one would have to put energy into it to bring it to rest, which is quite different from any particles which have actually been observed. The usual way of getting over this difficulty is to say that only the positive roots of equation [(6)] are to be taken into consideration and the others must simply be ignored.⁵²

In the classical theory, the energy of a particle varies continuously, but can never reach zero. (Its energy cannot drop below mc^2 .) So there is no possibility of a particle smoothly reaching a negative value, and the negative solutions of equation (6) can be ignored.

'This procedure is no longer permissible,' Dirac argued, 'when we use the quantum theory'. In the quantum theory, discrete energy jumps are possible, and the negative energy states are available. 'We must therefore look for some physical meaning for the states of negative energy'. Dirac noted that his particles with negative energy behaved in an electromagnetic field as if their charge was reversed. 'This immediately suggests a connection between electrons of negative energy and protons.

One might be tempted at first sight to say that an electron of negative energy IS a proton, but this, of course, will not do since protons certainly have positive energy and give up energy when brought to rest.

We must therefore establish the connection on a different basis. For this purpose we must take into consideration another property of electrons, namely the fact that they satisfy the Exclusion Principle of Pauli.

According to this principle it is impossible for two electrons ever to be in the same quantum state.

Now the quantum theory allows only a finite number of states for an electron in a given volume (if we put a restriction on the energy) so that if only one electron can go in each state there is room for only a finite number of electrons in a given volume. Thus we get the idea of a distribution of electrons. Let us now make the assumption that almost all the states of negative energy for an electron are occupied, so that the whole negative-energy domain is almost saturated with electrons. How will one of the few unoccupied negative-energy states now appear? It will be a sort of 'hole' in the distribution of negative energy electrons.⁵³

These holes would behave as particles with a positive charge *and* a positive energy (because it takes negative energy to fill the hole). 'It now appears reasonable to interpret this hole as a proton'.

Dirac was aware already in January 1930 that his theory met with a difficulty when confronted by standard electrodynamics.

The above theory of the relation between electrons and protons requires the existence everywhere of an infinite density of negative-energy electrons.

This infinite density would, according to Maxwell's equations applied in the ordinary way, give rise to an infinite electric field, and this has been put forward as an objection to the theory.

This objection can easily be met, however, by a re-interpretation of Maxwell's equations.

A perfect vacuum is to be considered as a region of space in which all the states of negative energy and none of those of positive energy are occupied.

The electron distribution in such a region must be assumed produce no field.

Only departures from this vacuum distribution must produce a field according to Maxwell's equations.⁵⁴

This physical situation was one in which there were no protons or electrons, only the infinite distribution of negative energy electrons. However a re-interpretation of Maxwell's equations did not solve all of Dirac's interpretative difficulties. How could an infinite distribution of matter be unobservable? And, from symmetry considerations, should not the mass of the electron and the mass of the hole be the same? The mass of the proton was known to be approximately 2000 times the mass of

⁵²Dirac Papers, Series II, Box 26, Folder 8.

⁵³Dirac Papers, Series II, Box 26, Folder 8.

⁵⁴Dirac Papers, Series II, Box 26, Folder 8.

the electron. Dirac's theory at this point did not include the interaction between the electrons and the holes, and he hoped that including the interaction terms would account for the mass difference.

Despite these difficulties, Dirac's work was received well by some in the physics community. The Science Service Washington DC Daily Science News bulletin reported on his talk under the headline 'New theory of space fills it with energy lack'. It reported that

Dr. Dirac's paper was praised highly by Sir Oliver Lodge who said: 'When developed this theory will probably lead us in the direction we wish to go'.

Sir Oliver also said that space is the important thing and that matter is merely incidental and accidental.⁵⁵

Sir Oliver Lodge was one of the most prominent British physicists of the previous era. He was perhaps seeing in Dirac's theory a rebirth of aether theories, which he had championed.⁵⁶ Dirac enjoyed the report enough to preserve it in his papers.

2.2. Correspondence networks

Dirac's correspondents expressed a mixture of admiration for, and doubts about, Dirac's new picture of the vacuum. Even before the January paper was published, Heisenberg wrote to Dirac expressing these emotions

Dear Dirac!

Many thanks for your letter. I heard about your new paper already a few days ago from Landau (via Gamow). I think I understand the idea of your new paper; it is certainly a great progress. But I cannot see yet, how the ratio of the [masses ?] etc. will come out. It seems to me already very doubtful, whether the terms of the electron (i.e. Sommerfeld formula) will not be completely changed by the interaction with the negative cells. One may hope, that all these difficulties will be solved by straight calculation of the interaction. In this case it would probably be necessary, to treat the interaction in a proper [sic] relativistic way, so Pauli and my paper might be of some use.⁵⁷

Here Heisenberg referenced his attempt with Wolfgang Pauli to establish a relativistic theory of their own.⁵⁸ Heisenberg illustrates how Dirac's ideas were circulated independently of Dirac himself. At this time, Ukrainian physicist George Gamow was moving between Bohr's Institute at Copenhagen and the Cavendish Laboratory at Cambridge. He kept up a correspondence with his friend from their Leningrad University days, Lev Landau.⁵⁹ Landau in turn corresponded with Heisenberg. In this paper, I chart only one topology of the rich transnational exchange of Dirac's ideas.

The traditional view has it that the work of mathematicians such as Weyl forced Dirac to abandon his interpretation of holes as protons because of the forced equivalence of electron and proton masses in the theory.⁶⁰ Dirac did cite Weyl, in addition to others, in print.⁶¹ However, Dirac was aware of the difficulty from his closest correspondents before Weyl's 1929 papers were published, and well before the second edition of his *Gruppentheorie und Quantenmechanik* appeared (1931).⁶² Hence, I differ from Tom Pashby's argument that the 'manifest relativistic invariance of

⁵⁵Dirac Papers, Series II, Box 26, Folder 8.

⁵⁶Bruce J Hunt, *The Maxwellians* (Ithaca, NY: Cornell University Press, 1991), appendix.

⁵⁷W. Heisenberg to P. Dirac, 7 December 1929, Dirac Papers, Series II, Box 1, Folder 9, 1929/7–12.

⁵⁸W. Heisenberg and W. Pauli, 'Zur Quantendynamik der Wellenfelder', *Zeitschrift Für Physik*, 56, nos. 1-2 (1929), 1–61; W. Heisenberg and W. Pauli, 'Zur Quantentheorie der Wellenfelder. II', *Zeitschrift Für Physik*, 59, nos. 3–4 (1930), 168–90.

⁵⁹Karl Hall, 'The Schooling of Lev Landau: The European Context of Postrevolutionary Soviet Theoretical Physics', *Osiris*, 23, no. 1 (2008), 230–59. doi:10.1086/591876

⁶⁰For example, Norwood Russell Hanson, 'Discovering the Positron (I)', *The British Journal for the Philosophy of Science*, 12, no. 47 (1961), 208; J. C. Polkinghorne, 'At the feet of Dirac', in *Reminiscences About a Great Physicist: Paul Adrien Maurice Dirac*, ed. by Behram N. Kursunoglu and Eugene Paul Wigner (Cambridge: Cambridge University Press, 1987), p. 221; Weyl's analysis is found in Hermann Weyl, *The Theory of Groups and Quantum Mechanics*, trans. by H. P. Robertson (London: Methuen & Co. Ltd, [1931] 1932).

⁶¹P. A. M. Dirac, 'Quantised Singularities in the Electromagnetic Field', *Proceedings of the Royal Society of London. Series A*, 133, no. 821 (1931), 60–72. doi:10.1098/rspa.1931.0130

⁶²Weyl's 1929 paper in *PNAS* was communicated 7 March 1929, his paper in *Zeitschrift für Physik* was communicated 9 May 1929.

Weyl's [1931] treatment of the interaction of the electron with the electromagnetic field had made it clear to Dirac that holes could not be protons'.⁶³ Already in his January paper he discussed this objection, probably raised in person in discussions at Cambridge in 1929. (Recall George Gamow's informal exposure to the theory, as reported by Heisenberg.) Rather than a stern correction to Dirac's beliefs, Weyl's analysis should be seen as an important contribution to working out Dirac's theory that fit within an already-ongoing conversation.

Aware of the conflicts between holes and protons, Dirac had hoped that including an account of the interaction between electrons and the holes/protons would somehow account for the mass difference. On 16 January 1930 Heisenberg wrote from Leipzig:

Dear Dirac:

Many thanks for the reprint of your new paper. I have calculated a little about the effect of interaction of the electrons in your theory. [...] One can prove, that electron and proton get the same mass.⁶⁴

This Dirac already knew,⁶⁵ however, Heisenberg had gone further: 'it is possible in principle to construct interactions, in which electron and proton get a different mass'. Dirac's intuition about including the interactions may have been correct. But the good news did not last long.

In this case, however, the ratio of the masses always get infinite, unless one goes entirely away from old interaction laws [...]. The Sommerfeld formula would be changed therefore. So I feel, that your theory goes very far away from any correspondence to classical laws; and also from experimental facts.⁶⁶

Pauli may have come to this conclusion by himself as well.⁶⁷ Other physicists quickly took up Dirac's work and its central questions. But they did not find results that met Dirac's hopes.

Igor Tamm, Dirac's friend and hiking companion, wrote on 3 March 1930 identifying another problem with Dirac's theory: the lifetime of an atom.⁶⁸ Simple atoms such as Hydrogen are remarkably stable. However, on Dirac's theory electrons and protons could annihilate one another (the electron filling the hole) and emit radiation.

I was very much interested in what you write about the probability of the annihilation processes, especially as I was just working out this probability myself when I got your letter. Considering a [?] positive-energy electron & a 'hole' in a volume V , I find by a calculation, closely analogous to that made for the scattering, that the probability of the annihilation taking place within a unit time equals to

$$Z = \frac{\pi d^2 c}{V} f(\beta) \quad (7)$$

where d stands for the classical diameter of the electron [...] & $f(\beta)$ is a function of the relative velocity $\beta = v/c$ of the electron & the hole: [...].

But the main difficulties are: 1) if one (tentatively & approximatively) applies the formula [(7)] to the case of bound electrons, one gets a ridiculously short value for the life-time of an atom, & 2) the frequency of the radiation emitted when an electron drops in a hole is of the order of magnitude of $\frac{mc^2}{h}$, where m is the mass of the electron & not of the proton, & that can't be reconciled with the existence of the cosmic radiation.⁶⁹

Tamm provides evidence that calculations within the 'hole' paradigm began quickly among Dirac's elite circle.⁷⁰ A narrow Kuhnian sense of 'paradigm' is appropriate here — in Tamm's letter we

⁶³Pashby, 'Dirac's Prediction of the Positron' (note 23), p. 453.

⁶⁴W. Heisenberg to P. Dirac, 16 January 1930, Dirac Papers, Series II, Box 2, Folder 1.

⁶⁵Dirac, 'A Theory of Electrons and Protons' (note 48), p. 360.

⁶⁶W. Heisenberg to P. Dirac, 16 January 1930, Dirac Papers Series II, Box 2, Folder 1.

⁶⁷See the letter from Tamm of 13 September 1930, Dirac Papers Series II, Box 2, Folder 3.

⁶⁸Igor Tamm to P. Dirac, 3 March 1930, Dirac Papers, Series II, Box 2, Folder 1. See also the correspondence given in A B Kojevnikov, ed., *Paul Dirac and Igor Tamm Correspondence; 1, 1928-1933*, MPI-Ph-93-80 (Munich: Max-Planck Institut für Physik, 1993).

⁶⁹Igor Tamm to P. Dirac, 3 March 1930, Dirac Papers, Series II, Box 2, Folder 1.

⁷⁰See the published work, Igor Tamm, 'Über die Wechselwirkung der freien Elektronen mit der Strahlung nach der Diracsehen Theorie des Elektrons und nach der Quantenelektrodynamik', *Zeitschrift Für Physik*, 62, nos. 7-8 (1930), 545-68. doi:10.1007/BF01339679

see a physicist working from Dirac's paradigmatic calculation within his overall scheme, the calculation of the scattering of an electron and a hole.⁷¹ This scattering calculation was included as §3 in Dirac's January paper. Here, Tamm takes up this paradigmatic example and tries to apply its structure to a new problem: the lifetime of an atom.

Vladimir Fock expressed different objections, again expressing a mixture of admiration and caution. Writing on 12 February, he began:

Dear Dr. Dirac!

I thank you very much for the copy of your last paper. I read it with the greatest interest, — but I cannot say that I was quite convinced by your theory.

Permit me to communicate you some doubts that raised in this connection.

Firstly: Can the Pauli exclusion principle be applied to a continuous set of states (with continuous eigenvalues)? I always thought that it can only be applied to an innumerable set of states (discrete eigenvalues), for in the formulation of this principle it seems to be necessary to numerate the states. For if the states may differ from one another by an infinitely small amount one can never say that all the states in a region of eigenvalues, however small, are filled up. In your theory, however, you apply the exclusion principle to all states, continuous as well as discontinuous.⁷²

Fock attacked perhaps the key move in Dirac's theory. The Pauli exclusion principle forbids electrons and protons from inhabiting the same state. It was this principle that allowed Dirac to build up his infinite distribution of negative energy electrons. These filled all the negative energy states, forcing some electrons to exist in states of positive energy. These positive energy electrons accounted for all actually observed electrons. This objection is important, and will be returned to below. Fock continued with a second objection:

Secondly: An infinite number of negative energy electrons per unit volume entails an infinite negative mass density all over the world. This must have astronomical consequences and may stand in conflict with the general theory of relativity.⁷³

Here Fock exhibited his holism, and wry understatement, worrying about how the different theories of physics could fit together.

Dirac replied to Fock on 21 February, though he did not keep a copy.⁷⁴ However, it is possible to glean something about Dirac's response from Fock's reply of 1 March.

I agree that the difficulties I mentioned may be, to some extent, removed on the lines given in your letter, but the necessity of introducing such things as impenetrable boxes (what are they made of? the unique really impenetrable 'box' is a finite universe!) and infinite but unobservable masses and charges seems to me a serious disadvantage of the theory. In every case, I will 'wait and see'.⁷⁵

Perhaps Fock's objections relating to the Pauli exclusion principle and continuous states led Dirac to consider quantizing space in infinitesimal, impenetrable boxes. This would create the conditions necessary for the discrete states of the negative energy electron distribution. Here we see physicists' ontological speculations circulating by private letter across the continent.

Fock and Tamm were not the only Soviet physicists interested in Dirac's new approach. Dirac and Dmitri Iwanenko kept up a correspondence during these years. In a letter dated 16 May 1930, Iwanenko connected Dirac's theory to a radical thought about the nature of space and time. Dirac's electron was still beset by the infinities of all relativistic quantum theories. This is usually expressed in terms of a high-frequency cut-off applied to integrals in the calculation. However, high-frequency

⁷¹This may also be an example of Cambridge style mathematical physics problem solving, as derived from examination problems, spreading out beyond England. See Andrew Warwick, *Masters of Theory: Cambridge and the Rise of Mathematical Physics* (Chicago, IL: University of Chicago Press, 2003).

⁷²V. Fock to P. Dirac, 12 February 1930, Dirac Papers, Series II, Box 2, Folder 1.

⁷³Ibid.

⁷⁴I have been informed that the letter has not been kept at Fock's archive in Russia. Personal communication, Alexei Kojevnikov, July 2015.

⁷⁵V. Fock to P. Dirac, 1 March 1930, Dirac Papers, Series II, Box 2, Folder 1.

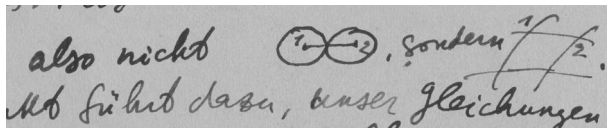


Figure 1. Detail from letter of D. Iwanenko to P.A.M. Dirac, 16 May 1930, Dirac Archives, Series II, Box 2, Folder 1. Florida State University Libraries, Tallahassee, FL.

also corresponds to short distance. In a scattering problem the high-frequency limit is the limit in which the distance between the two electrons reaches zero.

Suppose that space is discontinuous with $\Delta x = \frac{e^2}{mc^2}$. (here we have no choice because λ does not fit the electron; and the proton with m_p is out of consideration). Instead of saying that the electrons can not be closer together than the sum of their radii, we say that they are sitting in the space-nets' next possible point; so not [drawing] but [drawing]. (See figure 1.)⁷⁶

His next letter on 13 May announced his happiness about Dirac's upcoming visit to Russia that June. 'We expect the discussion with you about the point like space'. It is unknown whether Iwanenko and Fock discussed Dirac's letter to Fock expressing his conjecture that space was divided into impenetrable boxes. But it is likely, as Iwanenko and Fock were collaborators, and explored similar questions in print.⁷⁷ It may be that Iwanenko's thought was inspired by Dirac's correspondence with Fock.

Iwanenko's thoughts on space-quantization were not confined to his private correspondence. Writing with V. Ambarzumian in 1930, he explored the idea from within the perspective of Heisenberg and Pauli's electrodynamics, suggesting that the motion of electrons be confined to a three dimensional lattice structure.⁷⁸ Heisenberg, too, wrote on a theory of quantized space. Joan Bromberg writes that after his work with Pauli,

Heisenberg launched a new and radical attack on the difficulties he had enumerated; this work is recorded in letters he wrote to Bohr on 26 February and 10 March 1930. The most compelling motivation for Heisenberg's attack was the desire to solve the problem of the infinite self-energy of the electron. In analogy to classical theory, Heisenberg introduced the classical electron radius as a fundamental length. [...] Accordingly, Heisenberg proposed the construction of a 'lattice world' ('Gitterwelt') of cells of volume $(h/Mc)^3$, in which new relations would hold within the cells. He characterized this as the 'crudest method' by which a fundamental length could be introduced.⁷⁹

Unfortunately, Heisenberg realized, solving the self-energy problem would spoil relativistic invariance.⁸⁰

We can glean more about Dirac's thoughts along these lines from a letter from Heisenberg three months later on 14 July.

Dear Dirac:

Many thanks for your 'few lines' and the very clearer [sic] poetry about 'Quantization of space'. — I did not quite see the point in Iwanenko's paper, it seems to me, that it does not contain anything more, then I told in my Copenhagen lecture at [Easter ?]. So far the quantization of space is no theory at all, since it leaves all questions about relativistic invariance unanswered. [...] I think, there are two alternatives: Either there is

⁷⁶D. Iwanenko to P. Dirac, 16 May 1930, Dirac Papers, Series II, Box 2, Folder 1.

⁷⁷V. Fock and D. Iwanenko, 'Zur Quantengeometrie', *Physikalische Zeitschrift*, 30 (1929), 648; V. Fock and D. Iwanenko, 'Quantum Geometry', *Nature*, 123 (1929), 838.

⁷⁸V. Ambarzumian and D. Iwanenko, 'Zur Frage nach Vermeidung der unendlichen Selbstrückwirkung des Elektrons', *Zeitschrift Für Physik*, 64, nos. 7-8 (1930), 563-7; Helge Kragh and Bruno Carazza, 'From Time Atoms to Space-Time Quantization: The Idea of Discrete Time, ca 1925-1936', *Studies in History and Philosophy of Science Part A*, 25, no. 3 (1994), 457-8. doi:[http://dx.doi.org/10.1016/0039-3681\(94\)90061-2](http://dx.doi.org/10.1016/0039-3681(94)90061-2)

⁷⁹Joan Bromberg, 'The Impact of the Neutron: Bohr and Heisenberg', *Historical Studies in the Physical Sciences*, 3 (1971), 324.

⁸⁰I thank an anonymous referee for directing me to this literature. W. Heisenberg, 'The Self-Energy of the Electron', in *Early Quantum Electrodynamics: A Source Book*, ed. by Arthur Miller (Cambridge: Cambridge University Press, [1930] 1994), pp. 121-8; B. Carazza and H. Kragh, 'Heisenberg's lattice world: The 1930 theory sketch', *American Journal of Physics*, 63 (July 1995), 595. doi:[10.1119/1.17848](http://dx.doi.org/10.1119/1.17848)

no relativistic invariance, than there might exist a shortest wavelength and and [sic] a quantization of space might be the solution. Or it is always possible to have very short waves and very high energies and relativistic invariance. In the latter case it should be possible, to construct a theory for the electron and for radiation, where one puts $\mu = 0$ (electromagnetic = zero [sic]). For $\mu = 0$ it is not possible, to introduce a finite radius or anything like quantization of space.⁸¹

By μ Heisenberg was probably referring to the electromagnetic portion of the mass of the electron. A portion of the rest mass of the electron was thought to have originated in the energy associated with its charge.⁸² Heisenberg continued to write of their shared travels to Asia, and that he was practicing the game of go. He hoped to beat Dirac the next time they met. In a postscript Heisenberg concluded ‘Thinking nights and days about quantization of space’.⁸³ Dirac had introduced his picture of the vacuum as an infinite density of negative-energy electrons. His correspondence reveals a conversation stretching from Cambridge, through Leipzig, to Leningrad on possible radical consequences of Dirac’s theory for the nature of space-time.

What can we say of Weyl’s contribution to this debate? His papers in Spring 1929 were the first to suggest a *physical* interpretation for Dirac’s negative-energy states.⁸⁴ Weyl proposed that the negative-energy states could be protons. He may have been responding to the unease, expressed by Charles G. Darwin, that ‘To get a complete set [of solutions to the wave equation] we must double the number of solutions by admitting negative values of the energy, and we have at present little idea of what this means’.⁸⁵ Weyl’s work may have spurred Dirac to develop his own physical interpretation. Though Dirac’s 28 July 1928 letter to Klein indicates that the issue was on Dirac’s mind before Weyl’s publications, we do not have detailed records of Dirac’s thoughts at this date. The earliest surviving record of his thoughts is his letter to Bohr of 29 November 1929. Dirac differed with Weyl on the interpretation: negative-energy states were not protons, but rather *holes* in an infinite distribution of negative-energy electrons were to be identified as protons.⁸⁶ However, Weyl was *not* the first to point out that Dirac’s new interpretation ignored a crucial feature of the theory: the symmetry of positive- and negative-energy states, even in the case of interactions. (In fact, Weyl’s proton interpretation suffered the same defect.) Rather, it was Dirac’s correspondence network that was the first to register this objection, in the person of Heisenberg. This is reflected somewhat in Dirac’s published depiction of this exchange. In 1931, Dirac did not refer to his letters, but did cite Tamm’s published work, as well as Oppenheimer’s.⁸⁷ In retrospect, Dirac described Weyl as ‘the person who most definitely came out against’ the holes being protons.⁸⁸ This is a statement about the strength of Weyl’s mathematical analysis. It is not a statement of historical priority. The most important thing about Weyl’s involvement may be the observation that *he* not Dirac was the first to propose a physical interpretation of the negative-energy states required by Dirac’s electron theory. It was the Göttingen mathematician, not the Cambridge physicist, who propelled this line of metaphysical thinking.

2.3. The anti-electron

The next year, in October 1931, Dirac gave a series of lectures on quantum mechanics at the Institute for Advanced Study in Princeton, NJ. Notes were taken by B. Hoffman (probably Banesh Hoffman). These notes are interesting for three reasons. First they express Dirac’s view on the aesthetics of

⁸¹W. Heisenberg to P. Dirac, 14 July 1930, Dirac Papers, Series II, Box 2, Folder 3.

⁸²Kragh, *Quantum Generations* (note 6), pp. 105–8.

⁸³W. Heisenberg to P. Dirac, 14 July 1930, Dirac Papers, Series II, Box 2, Folder 3.

⁸⁴Weyl, ‘Elektron Und Gravitation. I’ (note 39); Weyl, ‘Gravitation and the Electron’ (note 39).

⁸⁵Darwin, ‘The Wave Equations of the Electron’ (note 21), p. 680; Donald Franklin Moyer, ‘Evaluations of Dirac’s Electron, 1928–1932’, *American Journal of Physics*, 49, no. 11 (1981), 1055–62.

⁸⁶In which Dirac cites Weyl’s (1929) *Zeitschrift für Physik* paper. Dirac, 1930, page 361.

⁸⁷Tamm, ‘Über die Wechselwirkung der freien Elektronen mit der Strahlung nach der Diracsehen Theorie des Elektrons und nach der Quantenelektrodynamik’ (note 70); Oppenheimer, ‘Note on the Theory of the Interaction of Field and Matter’ (note 51).

⁸⁸P. A. M. Dirac, ‘The Origin of Quantum Field Theory’, in *The Birth of Particle Physics*, ed. by Laurie M. Brown and Lillian Hoddeson (Cambridge: Cambridge University Press, 1983), p. 52.

approximation in physics (which will be relevant below). Second they express Dirac's view on the vacuum of photons — not just electrons (see also Bromberg).⁸⁹ Thirdly they indicate Dirac's new stage in his understanding of his equation and its consequences: he had abandoned the connections between electrons and protons, and now considered holes to be 'anti-electrons'. (This last point was famously made, in passing, in print the preceding May.⁹⁰) In proposing a new type of particle — against the dominant electron-proton model⁹¹ — we can read Dirac as choosing a simpler mathematics over a simpler ontology.

For the first point, it is notable that for all Dirac's paeans to mathematics throughout his career, he was not, and did not wish to be, a mathematician. He did not shy away from problems in which neat analytical solutions could not be found. A case in point arises in quantum theory when interactions between particles are taken into account. Recall that Dirac's theory of the electron described the motion of free electrons, or electrons in an external electromagnetic field. But what of interactions between the electrons themselves? Then the equations are not as neat, and solutions are in general impossible.

As soon as we have interaction between the particles of our system the problem becomes extremely difficult and in general involves equations that are so complicated that exact solutions cannot be found by present analytical methods; to attack these problems we must therefore make use of approximate methods; approximate methods do not necessarily have to be ugly and those we shall describe are based on general laws; all of physics is only approximation.⁹²

Here Dirac entertained the thought that aesthetics might preclude going any further in quantum theory. But he concluded by embracing the complications of the theory, asserting that there could be beauty in approximations (or if not beauty, at least not-ugliness). In fact, this was the whole of Dirac's discipline.

In a 1963 interview with Thomas Kuhn, Dirac elaborated on the aesthetics of approximation, and connected it to his early engineering training at Bristol.

Dirac: I think I owe a lot to my engineering training because it did teach me to tolerate approximations. Previously to that I thought any kind of an approximation was really intolerable. One should just concentrate on exact equations all the time. Then I got the idea that in the actual world all our equations are only approximate. We must just tend to greater and greater accuracy. In spite of the equation's being approximate they can be beautiful.⁹³

Kuhn tried to get Dirac to be more specific about what he meant by approximation. Was he referring to the mathematical sense in which a result may be only approximately close to a correct answer, or was he referring to an only-approximately-accurate *physical picture* underlying the calculations?

Kuhn: When you say approximate here, I think people often use the same term in two rather different ways. This may mean that they give approximate results in the sense that engineering equations do. It may mean that the whole physical theory with which the equations are involved is approximate in the sense that it's an approach to what's really there.

Dirac: Well, it's just neglecting a lot of factors. I meant it in that sense. The actual situation is far too complicated, you have to neglect a lot of factors.⁹⁴

⁸⁹Joan Bromberg, 'The Concept of Particle Creation Before and After Quantum Mechanics', *Historical Studies in the Physical Sciences*, 7 (1976), 186–7.

⁹⁰Dirac, 'Quantised Singularities in the Electromagnetic Field' (note 61).

⁹¹Kragh, *Quantum Generations* (note 6), pp. 174–7.

⁹²Dirac Papers, Series II, Box 26, Folder 15, 'Lecture on Quantum Mechanics', 72.

⁹³Interview with Dr. P. A. M. Dirac by Thomas S. Kuhn at Cambridge, England, May 6, 1963, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA, http://www.aip.org/history/ohilist/4575_2.html [accessed 21 October 2013].

⁹⁴Ibid.

Dirac replied in a way that did not provide a clear response to Kuhn's question. Were these neglected factors mathematical or physical, or did it make sense to differentiate them? Kuhn attempted to bring in a concrete example:

- Kuhn: Would you say that one gets the Bohr atom from quantum mechanics or that one gets to quantum mechanics from the Bohr atom by adding simply a factor that one has been ignoring previously, which would give a sense of approximation?
- Dirac: You mean by quantum mechanics the Schrödinger equation.
- Kuhn: The Schrödinger equation or the general — . I mean would that also be an example of an approximation in the sense that you have in mind?
- Dirac: Yes, it would be. Yes.
- Kuhn: Because clearly the logical relation of the earlier and later theory there is more complex; that is, it would be difficult to isolate a single — or group of extra things, that sort of added on
- Dirac: I think it's very likely that all our equations are only approximate. Our present quantum theory is probably only an approximation to the improvement of the future. I feel that everything might be an approximation and this comes very largely from the engineering training.⁹⁵

We see here that Kuhn was not totally successful in his attempts to get Dirac to specify what he meant by approximation. He set up the example of the progression from Bohr's quantum theory of the atom to Schrödinger's quantum mechanics, and Dirac agreed that this captures a sense of what Dirac meant by approximation. But it is still unclear whether this sense of approximation was at the level of adding mathematical terms, or at the level of changing physical pictures. Perhaps Dirac did not see the same distinction that Kuhn did. Perhaps he thought neglecting factors in one's approach to solving problems encompassed both the formal manipulations and the physical reconstruction.

Returning to Dirac's IAS lecture notes, and without going into the technical details, it is important to note Dirac's position on the vacuum of photons. While discussing the physics of symmetric wave functions he demonstrated the sometimes cavalier way in which physicists have made ontological statements. In his theory,

the total number of photons is not conserved since any atom may emit or absorb a photon; our theory of Einstein-Bose assemblies didn't apply to this case but the necessary modification is quite trivial; we merely postulate the existence of some new state, the zero state, such that when the photon is in the zero state it cannot be observed at all; if a photon is observed by an atom we say that it jumps into the zero state and if it is emitted we say that it jumps from the zero state into some other state; we assume that the number of photons in the zero state is infinite.⁹⁶

Much as the vacuum of electrons consisted of an infinite density of (negative-energy) electrons, Dirac's vacuum was populated by an infinite distribution of photons. This appears to have remained a constant of Dirac's thought since 1927.⁹⁷

At the end of his lectures at Princeton, Dirac turned briefly to his relativistic theory of the electron. He introduced the relativistic mass-energy relation with its energy-squared terms, and confronted his audience with the problem of what to do about the negative energy states. He briefly considered a proposal of Schrödinger's to modify the theory by neglecting the parts of the solutions to the wave function that describe transitions between positive and negative energy states. 'However, this method doesn't seem very satisfactory since the new wave equation is no longer relativistic; so probably this is

⁹⁵Ibid.

⁹⁶Dirac Papers, Series II, Box 26, Folder 15, 'Lecture on Quantum Mechanics', 102–3.

⁹⁷P. A. M. Dirac, 'The Quantum Theory of the Emission and Absorption of Radiation', *Proceedings of the Royal Society of London. Series A*, 114, no. 767 (1927), 260–1. doi:10.1098/rspa.1927.0039; Moyer, 'Origins of Dirac's Electron, 1925–1928' (note 18).

not really the correct line upon which one should proceed in order to resolve the difficulty of the negative energy levels'.⁹⁸ Another path given was Dirac's own, to imagine all the negative-energy states as occupied. 'This would be a formal way of getting over the difficulty but it does, of course, introduce the idea of an infinite number of electrons in the negative energy states [...]'.⁹⁹ This may have been a 'formal' solution, but Dirac was emphatic that it had physical content: 'These negative energy electrons are not to be considered as a mathematical fiction; it should be possible to detect them by experimental means'.¹⁰⁰ Nevertheless, it was not the negative-energy electrons themselves that Dirac focused on; rather it was the holes in their distribution. By 1931 however, these holes were no longer protons:

we shall refer to such a hole as an anti-electron; an anti-electron ought to have the same mass as an electron and this appears to be unavoidable; we should prefer to get a much larger mass so as to identify the anti-electrons with the protons, but this does not seem to work even if we take a Coulomb interaction into account.¹⁰¹

Here, and in the introduction to his published work on the magnetic monopole, was Dirac's prediction of the anti-electron.¹⁰²

One year later, in 1932, Dirac's new particle arrived.¹⁰³ Carl D. Anderson announced evidence for the existence of a particle with the same mass as the electron but opposite charge. He made no reference to Dirac's theory and christened the particle the positron. The next year Patrick Blackett and Giuseppe Occhialini confirmed its existence with explicit reference to Dirac.¹⁰⁴ Though we might say that the new particle was 'seen' before it was 'observed' by diverse experimentalists, and so point to the role of Dirac's theory in creating the conditions necessary for its discovery to be possible,¹⁰⁵ the results were largely seen at the time as confirming Dirac's prediction.¹⁰⁶ Subsequent historical analysis, however, has questioned the singularity of Blackett and Occhialini's result.¹⁰⁷ Dirac received congratulatory letters from Tamm and Iwanenko. However, Pauli was not impressed. He did not feel that the existence of the positron warranted accepting Dirac's ontological proposals about the vacuum. In May 1933 he wrote to Dirac: 'I do not believe in your perception of "holes" even if the "anti-electron" is proved'.¹⁰⁸

At the Solvay conference in 1933, Dirac made this connection between experiment and theory explicit. He opened his remarks with:

The recent discovery of the positively charged electron or positron has revived interest in an old theory about the states of negative kinetic energy of an electron, as the experimental results so far are in agreement with the predictions of the theory.¹⁰⁹

⁹⁸Dirac Papers, Series II, Box 26, Folder 15, 'Lecture on Quantum Mechanics', 131–2.

⁹⁹Dirac Papers, Series II, Box 26, Folder 15, 'Lecture on Quantum Mechanics', 132.

¹⁰⁰Dirac Papers, Series II, Box 26, Folder 15, 'Lecture on Quantum Mechanics', 132.

¹⁰¹Ibid., 134.

¹⁰²Dirac, 'Quantised Singularities in the Electromagnetic Field' (note 61).

¹⁰³Laurie M. Brown and Lillian Hoddeson, 'The Birth of Elementary Particle Physics: 1930–1950', in *The Birth of Particle Physics*, ed. by Laurie M. Brown and Lillian Hoddeson (Cambridge: Cambridge University Press, 1983), pp. 3–36; Carl D. Anderson and Herbert L. Anderson, 'Unraveling the Particle Content of Cosmic Rays', in *The Birth of Particle Physics*, ed. by Laurie M. Brown and Lillian Hoddeson (Cambridge: Cambridge University Press, 1983), pp. 131–54.

¹⁰⁴Carl D. Anderson, 'The Apparent Existence of Easily Deflectable Positives', *Science*, 76 (9 September 1932), 238–39; Carl D. Anderson, 'The Positive Electron', *Physical Review*, 43, no. 6 (Mar 1933), 491–4. doi:10.1103/PhysRev.43.491; Patrick M.S. Blackett and G.P.S. Occhialini, 'Some Photographs of the Tracks of Penetrating Radiation', *Proceedings of the Royal Society of London. Series A*, 139, no. 839 (1933), 699–726.

¹⁰⁵Hanson, 'Discovering the Positron (I)' (note 60); Norwood Russell Hanson, 'Discovering the Positron (II)', *The British Journal for the Philosophy of Science*, 12, no. 48 (1962), 299–313.

¹⁰⁶Behram N. Kursunoglu and Eugene Paul Wigner, eds., *Reminiscences About a Great Physicist: Paul Adrien Maurice Dirac* (Cambridge: Cambridge University Press, 1987).

¹⁰⁷Xavier Roqué, 'The Manufacture of the Positron', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 28, no. 1 (1997), 73–129.

¹⁰⁸Quoted in Pais, 'Playing with equations' (note 3), pp. 101–2.

¹⁰⁹Dirac Papers, Series II, Box 26, Folder 17. The published version is P. A. M. Dirac, 'Théorie du positron', in *Structure et propriétés des noyaux atomiques : rapports et discussions du septième Conseil de physique tenu à Bruxelles du 22 du 29 octobre 1933, sous les auspices de l'Institut international de physique Solvay* (1934) Paris: Gauthier-Villars, 203–12.

Dirac then pointed to two of his epistemic guides in evaluating physical theory: consistency with relativity, and self-consistency.

It has not been found to be possible to set up a relativistic quantum theory of the electron in which transitions from positive to negative energy values do not occur. It is thus no longer possible to assume that the energy is always positive without getting inconsistencies in the theory.¹¹⁰

For Dirac two facts forced a physicist to accept the reality of negative energy states: these states were required by relativity and self-consistency. Note that inter-theoretical consistency was also Dirac's guide for the initial formulation of his theory.¹¹¹ He had sought to create a relativistic theory of the electron that was consistent with the regular interpretation of quantum mechanics (the 'general transformation theory'). From this vantage point 'Two courses are now open to us.

Either we must find some physical meaning for the negative-energy states or we must say that the relativistic quantum theory of the electron is inaccurate to that extent to which it predict transitions from positive to negative-energy states.¹¹²

Dirac rejected this latter view. He considered an argument that quantum mechanics should not be applicable above certain energies. The idea of limits to the applicability of quantum mechanics at certain energy levels was common. Recall that since Lorentz, relativistic theories of the motion of the electron were bound up with theories of its inner structure.¹¹³ Dirac's theory posited a point electron — with no inner structure — but he did not think this was the final word on the matter. In fact, he wrote that it was

true that the present quantum mechanics cannot be expected to apply to phenomena in which distance of the order of the classical radius of the electron, $\frac{e^2}{mc^2}$, are important, since the present theory cannot give any account of the structure of the electron [...].

Pair creation phenomena involving the creation of positrons should only be possible starting at a photon energy of about twice the rest-mass of the electron, $2mc^2$. In order to compare the classical electron radius to the rest-mass energy of the electron, the electron radius must be expressed as an energy, leading to $\frac{hc}{e^2} \cdot mc^2$, which is much greater than the energy above mentioned [mc^2]. There was then no 'fundamental reason' why the quantum theory should not apply.¹¹⁴

'This result immediately suggests a connection between the negative-energy electronic states and the positron'. But as before, a straight identification of a positive particle with the negative energy states was impossible, because these particles did not have negative energy. Instead, relying on Pauli's exclusion principle, Dirac re-introduced his picture of holes in a uniform distribution of negative-energy electrons. 'Any unoccupied negative-energy state, though, being a break in the uniformity, we should expect to be observable as a sort of hole. We can now assume that such holes are the positrons'. Dirac also gave a physical interpretation to a common formal manoeuvre in relativistic theories: a cut-off. When calculating quantities such as the electron's mass or charge, physicists found that they had to integrate quantities over infinity. These integrals diverged to infinity. However, if you arbitrarily take some large number as the upper limit of your integration, you will not get an infinite divergence. Rather you will get some function of your cut-off. What could this procedure mean physically? Say your upper limit of integration was $137mc^2$. 'Physically, this means we assume that that part of the distribution of negative-energy electrons lying below an energy level of about $-137mc^2$ does not get polarized by the

¹¹⁰Ibid.

¹¹¹Here self-consistency refers to consistency within a given theory and inter-theoretical consistency refers to consistency between theories.

¹¹²Dirac Papers, Series II, Box 26, Folder 17.

¹¹³Olivier Darrigol, 'The Electron Theories of Larmor and Lorentz: A Comparative Study', *Historical Studies in the Physical and Biological Sciences*, 24, no. 2 (1994), 265–336, <http://www.jstor.org/stable/27757725>.

¹¹⁴Dirac papers, Series II, Box 26, Folder 17.

electric field in the way the theory gives'.¹¹⁵ At the end of the year, 12 December 1933, Dirac delivered his Nobel lecture on his theory.¹¹⁶

In pencil manuscript notes for a 'Theory of the Positron' in 1934 Dirac put forth the most well-known metaphor for his 'uniform distribution of negative energy electrons': with shades of Thales, the vacuum was a sea. After beginning in a similar manner to his Princeton lectures, Dirac introduced his negative-energy electrons and the application of the Pauli exclusion principle. (The seriousness of Fock's earlier objection to the use of the principle for the case of continuous eigenstates seems to have faded.)

We can then assume that in the world as we know it, nearly all the $-ve$ energy states are filled, with one electron in each state, and that the vacancies or holes in the $-ve$ energy distribution are the positrons.

A rough picture is to compare it with the sea, a bubble in the sea corresponding to a proton, just as a drop of water in the air corresponds to an ordinary electron.

The sea is a bottomless one, and we do not consider at all how it is supported from below, but are interested only in events near its surface.¹¹⁷

Dirac included sketches in his working papers of the sea and its holes (figure 2). On the back of this page, Dirac made geometric sketches, perhaps indicating that he was using his 'hidden' geometry to work out his theories of the positron in the mid-1930s (see below).¹¹⁸ He also speculated on the existence of an anti-particle partner to the proton, to preserve symmetry: 'Probability of $-ve$ proton, to preserve symmetry between $+ve$ & $-ve$ charge'.¹¹⁹

Moving past unpublished writings, a standard reference among English physicists was to Dirac's 1934 article in the *Mathematical Proceedings of the Cambridge Philosophical Society*, as for example in Homi Bhabha's work.¹²⁰ Though Heisenberg would point to Dirac's 1930 *Principles of Quantum Mechanics* as the source of his 'positron' theory in 1934, Dirac's theory was alive enough in the community by 1936 that Heisenberg and H. Euler could write on the 'Consequences of Dirac's Theory of the Positron' without citing Dirac at all.¹²¹ Dirac's semi-formal lecture notes were also circulated in the physics community. A Worldcat search finds 16 surviving copies of Dirac's 1934–35 IAS Princeton lectures in libraries, which were mimeographed and distributed. These notes contained a brief concluding section of the 'Theory of the Positron' showing how to address some infinities in the theory.¹²² Dirac's picture of the positron as a hole in a sea of negative-energy electrons remained the paradigm for analyzing relativistic particles throughout the 1930s, as evinced for example by the work of Weisskopf. More details are given in Mehra and Rechenberg and by Moyer.¹²³

¹¹⁵Dirac papers, Series II, Box 26, Folder 17.

¹¹⁶P. A. M. Dirac, 'Nobel Lecture: Theory of Electrons and Positrons', *Nobel Media AB*, [1933] 2013, http://www.nobelprize.org/nobel_prizes/physics/laureates/1933/dirac-lecture.html [accessed 21 October 2013].

¹¹⁷Dirac Papers, Series II, Box 36, Folder 22.

¹¹⁸See also Dirac's 6 May 1967 interview with T.S. Kuhn, available at http://www.aip.org/history/ohilist/4575_2.html [accessed 21 October 2013]; Peter Galison, 'The Suppressed Drawing: Paul Dirac's Hidden Geometry', *Representations*, no. 72 (2000), 145–66.

¹¹⁹Dirac Papers, Series II, Box 36, Folder 22.

¹²⁰P. A. M. Dirac, 'Discussion of the Infinite Distribution of Electrons in the Theory of the Positron', *Mathematical Proceedings of the Cambridge Philosophical Society*, 30, no. 2 (March 1934), 150–63. doi:10.1017/S030500410001656X; H. J. Bhabha, 'The Scattering of Positrons by Electrons with Exchange on Dirac's Theory of the Positron', *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 154, no. 881 (1936), 195–206.

¹²¹P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford: Clarendon Press, 1930); W. Heisenberg, 'Bemerkungen Zur Diracschen Theorie Des Positrons', *Zeitschrift Für Physik*, 90, nos. 3–4 (1934), 209–31. doi:10.1007/BF01333516; W. Heisenberg and H. Euler, 'Folgerungen aus der Diracschen Theorie des Positrons', *Zeitschrift Für Physik*, 98, nos. 11–12 (1936), 714–32; W. Heisenberg and H. Euler, 'Consequences of Dirac Theory of the Positron', trans. by W. Korolevski and H. Kleinert, *ArXiv Physics E-Prints*, May [1936] 2006.

¹²²P. A. M. Dirac, *Lectures on Quantum Electrodynamics* (Princeton, NJ: Institute for Advanced Study, 1935), Chapter IV.

¹²³V. Weisskopf, 'Über Die Selbstenergie Des Elektrons', *Zeitschrift Für Physik*, 89, nos. 1–2 (1934), 27–39; V. Weisskopf, 'Über die Elektrodynamik des Vakuums auf Grund der Quantentheorie des Elektrons', *Det Kgl. Danske Videnskaberne Selskab. Matematisk-Fysiske Meddelelser*, XIV, no. 6 (1936), 1–39; V. F. Weisskopf, 'On the Self-Energy and the Electromagnetic Field of the Electron', *Physical Review*, 56 (July 1939), 72–85. Mehra and Rechenberg, *The Completion of Quantum Mechanics 1926–1941 Part IV*, 3; Moyer, 'Evaluations of Dirac's Electron, 1928–1932' (note 85).

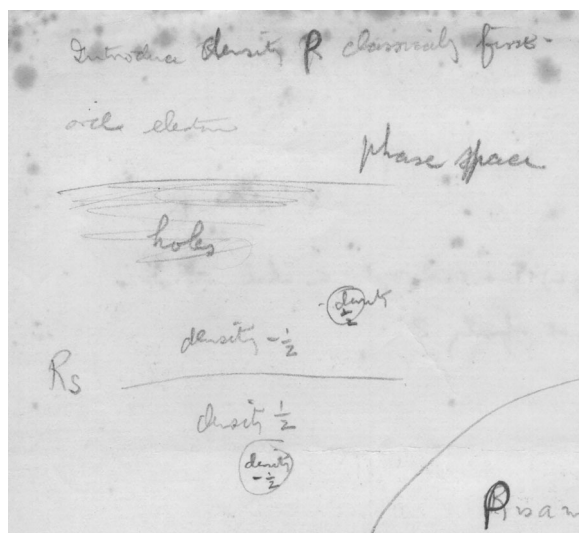


Figure 2. Dirac's pencil sketch, top left, of the sea of electrons and negative-energy electrons and holes. The diagram is marked "ord[inary] electrons" and "holes." Dirac Papers, Series II, Box 36, Folder 22. Florida State University Libraries, Tallahassee, FL.

2.4. A connected Dirac

This account foregrounds the evolving metaphysical positions surrounding Dirac's theory in the 1920s and 30s. As has been shown through an analysis of private correspondence and published papers, Dirac's theory was tied up with evolving conceptions of space and time. This period was an inflection point in the trajectory of our understanding of space, time, and the vacuum. It was not a controversy in the Sociology of Scientific Knowledge sense.¹²⁴ However it reveals an opening, a moment of indecision and contingency, a point at which our fundamental concepts — had things gone slightly differently — may have been altered.

Dirac began in 1928 with formal manoeuvres in interpreting his equation. At this time workers in relativistic quantum theory were used to simply arbitrarily discarding unpleasant terms in their equations. (This was to be derided later on as 'subtraction physics' by Pauli and others¹²⁵) Dirac put forth the option of simply ignoring the $-ve$ energy solutions to his wave equation. In the minds of Dirac's most prominent correspondents, his theory was from the outset intimately related to space and time. Recall that in February 1928 at Copenhagen, Niels Bohr related Dirac's work to Oskar Klein's five-dimensional theory of electromagnetism and gravity. However, the negative energy solutions could not be ignored, a fact Dirac discovered — perhaps in consultation with Werner Heisenberg — still in 1928. That summer in Leipzig he gave a talk suggesting that his theory would have to be only approximately true. More than that, he thought a thoroughgoing solution to the problem of the existence of negative-energy states would require much more conceptual work; in Leipzig these were connected briefly to ideas about the difference between past and future.

Hermann Weyl was the first to publish a physical interpretation of Dirac's negative-energy states in 1929. Later that year we know, again from his correspondence with Bohr, that Dirac was working on a physical interpretation of his own. On 26 November Dirac let loose his picture of 'holes' in an infinite sea of negative-energy electrons within the physics community. On the first day of the New Year in 1930 his interpretation was published in the *Proceedings of the Royal Society of London*. Here Dirac reconfigured our conceptions of the nature of empty space. Rather than a truly empty arena for

¹²⁴Sophia Roosth and Susan Silbey, 'Science and Technology Studies: From Controversies to Posthumanist Social Theory', in *The New Blackwell Companion to Social Theory*, ed. Bryan S. Turner (Malden, MA: Blackwell Publishing Ltd., 2009), 451–73.

¹²⁵Schweber, *QED and the men who made it* (note 9), pp. 76–152.

physics — or even a dynamic manifold in a relativistic picture — the vacuum was full. In each piece of ‘empty’ space was an infinite distribution of negative-energy particles. Though only the holes in this sea would be observable.

By February, Vladimir Fock raised an objection to a fundamental plank of Dirac’s theory: the application of Pauli’s exclusion principle. In order for the infinite sea of negative-energy electrons to be ‘full’ the exclusion principle needed to be in place. Otherwise the electrons, like photons, would all collapse into the same (or nearly the same) state. There would be no definite ‘holes’ with particle-like properties. Dirac replied by letter to Fock, and kept a copy of the Soviet’s March reply. Dirac appears to have proposed that empty space be divided into impenetrable boxes, so as to provide the boundary conditions necessary for the application of Pauli’s principle. By May another Soviet physicist had engaged with Dirac’s theory: Dmitry Iwanenko. He proposed that the relativistic cut-offs used in the theory could find a physical manifestation in a quantization of space. Dirac then discussed these points with Heisenberg, who was so animated by the possibility that he included a postscript in his letter that he was ‘Thinking nights and days about quantization of space’. During these two years from 1928 to 1930 the foundational notions of space, time, and vacuum were at play in physics.

It is worth pausing here to emphasize the knowledge networks Dirac was enmeshed in. Though he is often represented as ‘strange’ and reserved,¹²⁶ Dirac was an active member of the scientific community. He kept up a detailed transnational correspondence. He circumnavigated the globe.¹²⁷ He published in a variety of journals. He was part of an ‘epistemic culture,’ or (sub-)‘thought collective’.¹²⁸ It is to the details of Dirac’s expression of this epistemic culture that I now turn.

3. Mathematical beauty

So far, this paper has detailed the shifting ontological commitments that were a part of Dirac’s relativistic theory of the electron. What about epistemological strategies? This section will focus more on Dirac than on his correspondence network, though undoubtedly aspects of his thinking were shared across the community of physicists. According to the contemporary report of Charles G. Darwin, referring to Dirac’s 1928 theory, ‘Dirac’s guiding principle [was] that the ‘Hamiltonian equation’ must be linear’.¹²⁹ Was Darwin right? Was Dirac’s guiding principle so formal, and so limited? Let us see if we can move from this linearization to something deeper and more general. According to Dirac’s published papers, the reason the the Hamiltonian must be linear is that linearity preserved a sense of causality: ‘so that the wave function at any time determines the wave function at any later time’. Dirac continued with a second reason. ‘The wave equation of the relativity theory must also be linear in W if the general interpretation is to be possible’.¹³⁰ Causality and consistency with Dirac’s ‘darling’ general transformation theory were the reasons undergirding the formal requirement of linearity.

We can further inquire as to why causality and consistency with the general theory were important. For this we may stay within a close reading of Dirac’s 1928 introduction of his theory. Dirac started the paper with the observation that present theories of quantum mechanics — which included the presupposition that the electron was a point-like object — were not empirically adequate. They could not account for ‘duplexity’ phenomena that Goudsmit and Uhlenbeck had attributed to the electron having spin angular momentum in 1926 (with Bohr’s postscript).¹³¹ But there

¹²⁶Graham Farmelo, *The Strangest Man: The Hidden Life of Paul Dirac, Mystic of the Atom* (New York: Basic Books/Perseus Books, 2009).

¹²⁷Kragh, *Dirac* (note 10), Ch. 4.

¹²⁸Ludwik Fleck, *Genesis and Development of a Scientific Fact* (Chicago: The University of Chicago Press, [1935] 1981); K Knorr-Cetina, *Epistemic Cultures: How the Sciences Make Knowledge* (Cambridge, MA: Harvard University Press, 1999).

¹²⁹Darwin, ‘The Wave Equations of the Electron’ (note 21), p. 655.

¹³⁰Dirac, ‘The Quantum Theory of the Electron’ (note 15), p. 612.

¹³¹E. Uhlenbeck and S. Goudsmit, *Naturwissenschaften* 13, no. 47 (1925), 953; G. E. Uhlenbeck and S. Goudsmit, ‘Spinning Electrons and the Structure of Spectra’, *Nature*, 117 (February 1926), 264–5 and postscript by N. Bohr; S. A. Goudsmit, ‘The Discovery of the

were challenges to the picture of a spinning electron, that had been known since Lorentz: given known limits on the radius of the electron and the required angular momentum, each electron would have to spin faster than the speed of light.¹³² ‘The question remains as to why Nature should have chosen this particular model for the electron instead of being satisfied with the point-charge’.¹³³ Dirac continued with a statement about the incompleteness of the existing theories. If there were some incompleteness, and it could be rectified, if the theory could be made whole, perhaps one would not need recourse to the spinning electron model. One could continue with the point charge picture.

One would like to find some incompleteness in the previous methods of applying quantum mechanics to the point-charge electron such that, when removed, the whole of the duplexity phenomena follow without arbitrary assumptions. In the present paper it is shown that this is the case, the incompleteness of the previous theories lying in their disagreement with relativity, or, alternatatively [*sic*], with the general transformation theory of quantum mechanics.¹³⁴

Here Dirac emphasized the importance of consistency, of seeing physical theory as part of a whole. It is impossible to parse out whether Dirac thought consistency with relativity or with his transformation theory was more important, or whether he thought one entailed the other (no drafts of the paper were preserved).

We can see here an indication of an aesthetics of physics: proper theories should fit together. The theory of the electron should be relativistic, should fit with the rest of quantum mechanics, and should fit with Bohr’s correspondence principle. It must be noted, however, that these other theoretical desiderata are not aesthetic in nature.^{135,136} That is to say, the stipulation that a theory should fit with the correspondence principle was part of an aesthetics, but the details of the correspondence principle itself were not aesthetic considerations. Simplicity also came into play: ‘It appears that the simplest Hamiltonian for a point-charge electron satisfying the requirements of both relativity and the general transformation theory leads to an explanation of all duplexity phenomena without further assumption’.¹³⁷

Moving out from Dirac’s published paper, what can we say of his methodology, his strategies for understanding the world? According to McAllister, Dirac’s later reconstructions of his methodology relied on aesthetics extremely strongly. Rather than playing a role only in the context of discovery, Dirac held that aesthetic considerations came into play in the context of justification.¹³⁸ According to Pais, Dirac’s method was ‘playing’ with equations.¹³⁹ This Pais drew from his personal interactions with Dirac at the Institute for Advanced Study in Princeton, NJ, in the 1940s and 50s.

By far, the most revealing insight I gained during those discussions concerned the Dirac way of playing with equations, which can be summed up like this: First play with pretty mathematics for its own sake, then see whether this leads to new physics.

Throughout most of his life, that attitude is manifest in his writings.¹⁴⁰

Electron Spin’, in *Foundations of Modern EPR*, ed. by G.R. Eaton, S.S. Eaton, and K.M. Salikhov, trans. by J. H. van der Waals (Singapore: World Scientific, [1971] 1998), [online].

¹³²Nadia Robotti, ‘Quantum Numbers and Electron Spin: The History of a Discovery’, *Archives Internationales d’Histoire Des Sciences*, 40, no. 125 (1990), 305–31.

¹³³Dirac, ‘The Quantum Theory of the Electron’ (note 15), p. 610.

¹³⁴*Ibid.*, p. 610.

¹³⁵Dirac’s understanding of Bohr’s correspondence principle in terms of taking a large-quantum-number limit of a quantum equation dates from Fowler’s lectures at Cambridge in the mid-1920s. See Dirac Papers, Series II, Box 32, Folder 4, ‘Fowler’s Lectures of Quantum Theory. Cambridge 1924–1926’, [15].

¹³⁶‘Our problem is to obtain a wave equation of the [linear] form [. . .] which shall be invariant under a Lorentz transformation and shall be equivalent to [the classical case] in the limit of large quantum numbers’. Dirac, ‘The Quantum Theory of the Electron’ (note 15), p. 613.

¹³⁷*Ibid.*, p. 610.

¹³⁸J.W. McAllister, ‘Dirac and the Aesthetic Evaluation of Theories’, *Methodology and Science*, 23 (1990), 87–102.

¹³⁹Pais, ‘Playing with equations’ (note 3), p. 98.

¹⁴⁰*Ibid.*, p. 109.

Pais quotes the methodological reflections in Dirac's 1931 article on magnetic monopoles,¹⁴¹ an essay on 'The Relation Between Mathematics and Physics' (of which more below),¹⁴² and an interview with Kuhn on 7 May 1963: 'I think its just a peculiarity of myself that I like to play about with equations, just looking for beautiful mathematical relations which maybe don't have any physical meaning at all. Sometimes they do'.¹⁴³

3.1. Beauty in action

Pais is not the only author to have noted Dirac's writing on mathematical beauty. Olivier Darrigol is another.¹⁴⁴ But Kragh has the most nuanced discussion.¹⁴⁵ Chronology is important here. For though an emergent aesthetics can be gleaned from Dirac's 1928 work, there was no mention of beauty there. In fact a full-throated expression of Dirac's views on mathematical beauty would not come until 1939. It is worth quoting the concluding paragraph of Kragh's biography:

If a conclusion is to be drawn from Dirac's own career with respect to the scientific value of the principle of mathematical beauty, it appears to me to be the following. Many of his most important results were products of his belief in the power of mathematical reasoning; however, the principle of mathematical beauty, in its more elaborate meaning, proved to be a failure in Dirac's career. He applied it in particular in his persistent attempts to formulate an alternative quantum electrodynamics, and these attempts, as far as we can tell, were failures. In Dirac's scientific life, the mid-1930s marked a major line of division: all of his great discoveries were made before that period, and after 1935 he largely failed to produce physics of lasting value. It is not irrelevant to point out that the principle of mathematical beauty governed his thinking only during the later period.¹⁴⁶

According to Kragh, then, Dirac definitely espoused and used a principle of mathematical beauty, but perhaps all for the worst. My primary intervention in this debate will be the discussion of previously under-analyzed lecture notes Dirac gave in 1927, expressing his views on mathematical beauty during his most productive period.¹⁴⁷ This challenges any 'major line of division' that separates Dirac's commitment to mathematical beauty throughout his career.

Let us first of all then return to Pais's analysis. He qualifies his remarks with 'throughout most of [Dirac's] life,' but cites a paper as early as 1931 as evidence for Dirac's playing around with pretty mathematics. Dirac's 1931 paper is notable in the history of mathematical physics for being an early and important application of the ideas of global topology to modern physics — Dirac showed how, if there were to exist one magnetic monopole in the universe, it would explain why charge comes in discrete units. This result lay dormant until later in the century when experimentalists began actually 'hunting' for monopoles.¹⁴⁸ By the 1960s topological methods became increasingly important across theoretical physics.¹⁴⁹ The paper is also remarkable for its statement of Dirac's prediction of the anti-electron. In addition to all this, it begins with reflections on scientific methodology:

The steady progress of physics requires for its theoretical formulation a mathematics that gets continually more advanced. This is only natural and to be expected. What, however, was not expected by the scientific workers of the last century was the particular form that the line of advancement of the mathematics would take, namely, it was expected that the mathematics would get more and more complicated, but would rest on a permanent basis

¹⁴¹Dirac, 'Quantised Singularities in the Electromagnetic Field' (note 61).

¹⁴²P. A. M. Dirac, 'The Relation between Mathematics and Physics', *Proceedings of the Royal Society (Edinburgh)*, 59, no. Part II (1938–39), 122–9.

¹⁴³Quoted in Pais, 'Playing with equations', p. 110.

¹⁴⁴Darrigol, *From c-Numbers to Q-Numbers* (note 2), p. 302.

¹⁴⁵Kragh, *Dirac* (note 10), pp. 275–92.

¹⁴⁶*Ibid.*, p. 292.

¹⁴⁷But see Darrigol, *From c-Numbers to Q-Numbers* (note 2), p. 345.

¹⁴⁸Andrew Pickering, 'Constraints on Controversy: The Case of the Magnetic Monopole', *Social Studies of Science*, 11, no. 1 (February 1981), 63–93.

¹⁴⁹Aaron Sidney Wright, 'The Advantages of Bringing Infinity to a Finite Place: Penrose diagrams as objects of intuition', *Historical Studies in the Natural Sciences*, 44, no. 2 (2014), 99–139. doi:10.1525/hsns.2014.44.2.99

of axioms and definitions, while actually the modern physical developments have required a mathematics that continually shifts its foundations and gets more abstract.¹⁵⁰

Dirac may have been thinking of David Hilbert's famous problems for mathematicians, introduced in 1900. The sixth problem (as published in 1902) was the axiomatization of physics.¹⁵¹ As examples, Dirac considered

Non-euclidean geometry and non-commutative algebra, which were at one time considered to be purely fictions of the mind and pastimes for logical thinkers, have now been found to be very necessary for the description of general facts of the physical world. It seems likely that this process of increasing abstraction will continue in the future and that advance in physics is to be associated with a continual modification and generalisation of the axioms at the base of the mathematics rather than with a logical development of any one mathematical scheme on a fixed foundation.¹⁵²

Here Dirac expressed the opinion that Hilbert's programme would be fruitless — rather than searching for a stable axiomatization for physics, one should be constantly modifying and generalizing the axioms upon which physics was based. He characterized this as a process of increasing abstraction.

In his 1931 paper Dirac presented his 1930 work on the theory of electrons and protons as a paradigm for this methodology.¹⁵³ In his 'Lectures on Quantum Mechanics' in October 1931 in Princeton he expressed the view that the entirety of modern quantum mechanics might be seen this way.

Formerly we used to think that the formalism of the Newtonian theory would, if pushed far enough, give all the results we wanted; but now we are introducing new formalisms — new ideas — into the scheme of physics; we are changing our axioms.¹⁵⁴

Dirac continued his reflection of the methodology of physics with advice for his fellow researchers.

The most powerful method of advance that can be suggested at present is to employ all the resources of pure mathematics in attempts to perfect and generalise the mathematical formalism that forms the existing basis of theoretical physics, and after each success in this direction, to try to interpret the new mathematical features in terms of physical entities (by a process like Eddington's Principle of Identification).¹⁵⁵

Here we have the suggested order of Dirac's own thinking, as expressed by Pais: first make mathematical inroads, then search for physical meaning.¹⁵⁶ What to make of the reference to Eddington? According to Darrigol, Eddington was an idealist, believing in the primacy of the mind in determining the reality of existence. Kragh calls this rationalism.¹⁵⁷ Narrowly, Eddington conceived of his principle as applying only to tensors, thus Dirac's usage was already an expansion of its narrow meaning.¹⁵⁸ Eddington expressed these views in books on the general theory of relativity in the early 1920s.¹⁵⁹ Within this broader philosophical outlook, Eddington's 'principle of identification' was a methodological dictum. '[T]he mathematics of a physical theory had to be developed at an a priori level before the identification of physically accessible quantities took place'.¹⁶⁰ 'Identification' here means just that: pointing to a mathematical expression and asserting that it represented the

¹⁵⁰Dirac, 'Quantised Singularities in the Electromagnetic Field' (note 61), p. 60.

¹⁵¹David Hilbert, 'Mathematical problems', *Bulletin of the American Mathematical Society*, 8 (10) ([1900] 1902), 437–79.

¹⁵²Dirac, 'Quantised Singularities in the Electromagnetic Field' (note 61), p. 60.

¹⁵³A recent paper by the author may possibly be regarded as a small step according to this general scheme of advance' (*ibid.*, p. 61).

¹⁵⁴Dirac Papers, Series II, Box 26, Folder 15, 'Lecture on Quantum Mechanics', p. 1.

¹⁵⁵*Ibid.*, p. 60.

¹⁵⁶See also Darrigol, *From c-Numbers to Q-Numbers* (note 2), p. 308.

¹⁵⁷Helge Kragh, 'Cosmo-Physics in the Thirties: Towards a History of Dirac Cosmology', *Historical Studies in the Physical Sciences*, 13 (1982), 69–108.

¹⁵⁸I thank an anonymous referee for making this point. See Thomas Ryckman, *The Reign of Relativity: Philosophy in Physics 1915–1925* (New York: Oxford University Press, 2005), p. 205.

¹⁵⁹Arthur Stanley Eddington, *Space, Time and Gravitation: An Outline of the General Relativity Theory* (Cambridge: Cambridge University Press, 1920); Arthur Stanley Eddington, *The Mathematical Theory of Relativity* (Cambridge: Cambridge University Press, 1923).

¹⁶⁰Darrigol, *From c-Numbers to Q-Numbers* (note 2), p. 300.

world. According to Kilmister, this was a form of conventionalism.¹⁶¹ The kernel of these labels is the prominence of the role of human cognition or agreement in determining our picture of physical reality.

Was Dirac Eddingtonian? Darrigol suggests that he was. As the earliest evidence for this perspective he cites an undated manuscript, probably from 1924, in which Dirac used the word ‘beauty’ parenthetically:

The modern physicist does not regard the equations he has to deal with as being arbitrarily chosen by nature [...]. In the case of gravitational theory, for instance, the inverse square law of force is of no more interest — (beauty)? — to the pure mathematician than any other inverse power of distance. But the new law of gravitation has a special property, namely its invariance under any coordinate transformation, and being the only simple law with this property it can claim attention from the pure mathematician.¹⁶²

Darrigol’s emphasis in analyzing this quote is on the first sentence, claiming that it shows Dirac shared Eddington’s belief in the necessity of physical law, with some qualification: ‘it is not the search of the mind for permanence but its predilection for mathematical beauty which enforces the necessity of the laws’.¹⁶³ But while this is surely an example of Dirac’s use of the word, what was he really saying here about beauty? That it may sit alongside interest for pure mathematicians. Here Dirac was motivating his presentation of relativity to a group of mathematicians.¹⁶⁴ The gravitational law of Newton may be of no more interest/beauty to mathematicians than any other law of the same form, but Einstein’s gravitational theory was different (it was invariant under any coordinate transformation). Thus, while this is undoubtedly Dirac writing on the theme of beauty, it seems to me that there is less reason for ascribing Eddington’s metaphysical idealism to Dirac. This is especially true if the dating of this document to 1924 is correct — this would be just at the time of Dirac’s exposure to Bohr’s correspondence principle and prior to his exposure to Heisenberg’s positivist writing.¹⁶⁵ In support of an Eddingtonian Dirac, we are then left with Dirac’s methodological statement of 1931 to follow ‘a process like’ Eddington’s methodological principle. This was hardly a full-throated endorsement of Eddington’s metaphysics. The view expressed in this quote was not a rationalism but an aesthetics; not the view that the human mind determined the form of natural laws, but the view that nature’s laws could be identified by their aesthetic qualities.

Moving from Darrigol’s discussion to Kragh’s, the evidence that Dirac articulated a principle of mathematical beauty ‘in its more elaborate meaning’ prior to Kragh’s mid-1930s delineation comes from unpublished ‘Lectures on Modern Quantum Mechanics’. These notes are undated, but the Florida State University archivists, the keepers of Dirac’s papers, have provided the note that these are the ‘First lecture course of quantum mechanics; probably beginning in Oct 1927’. This note may have been based on information given by Dirac himself.¹⁶⁶

Dirac began this course with reflections on the empirical inadequacy of classical theories of physics.

It has long been known that classical electrodynamics meets with difficulties in the explanation of atomic phenomena. These difficulties are of a very fundamental nature and appear to make the classical theory irreconcilable with the facts, and has led to the introduction of a quantum theory differing from the classical theory on certain fundamental points.¹⁶⁷

It should be noted that he starts from experience here, not from an underlying mathematical structure, as Eddington might have. But he quickly moves on to aesthetics:

¹⁶¹C. W Kilmister, *Eddington’s Search for a Fundamental Theory: A Key to the Universe* (Cambridge: Cambridge University Press, 1994), p. 61.

¹⁶²Darrigol, *From c-Numbers to Q-Numbers* (note 2), pp. 30102.

¹⁶³*Ibid.*, p. 302.

¹⁶⁴Henry Frederick Baker’s Cambridge tea parties, see *ibid.*, pp. 295–6.

¹⁶⁵Mara Beller, *Quantum Dialogue: The Making of a Revolution* (Chicago, IL: University of Chicago Press, 1999), Ch. 4.

¹⁶⁶These notes are also available in the Archive for the History of Quantum Physics, American Philosophical Society, Microfilm 36, document 11.

¹⁶⁷Dirac Papers, Series II, Box 26, Folder 2, ‘Lectures on Modern Quantum Mechanics’, 1.

The chief thing that the quantum theory has had to fight against is the fact that classical electrodynamics is a very beautiful elegant complete self-consistent theory based on a few simple assumptions, and it appears that any modification of such a theory ~~should not be true~~ must be ugly, arbitrary, and not what we would expect nature to be like.¹⁶⁸

The beauty of classical electrodynamics has been the *chief* obstacle of the quantum theory. Here Dirac began with beauty but crossed it out in favour of other aesthetic concerns: it was elegant, it was based on simple assumptions. (This polysemy reinforces the necessity for this analysis of maintaining Dirac's own diffuse understanding of beauty.) The classical theory had related mathematical qualities: it was complete and self-consistent. Changing such a theory would make it ugly and arbitrary, which was 'not what we would expect nature to be like'. Here Dirac's aesthetic expectations of nature are clearly expressed.

However, as his students must have known, the quantum theory would overcome in this agonistic encounter.

This objection is no longer valid since the quantum theory, after passing through many stages and having its ~~final~~ fundamental concepts changed more than once, has now reached a ~~stage~~ form in which although not quite complete, is based on simple generalizations, and is as beautiful, and in certain respects more beautiful than the classical theory. This has been brought about by the fact that the new Q.T. required very few changes from the classical theory, these changes being of a fundamental nature, so that many of the features of the classical theory to which it owes its attractiveness can be taken over directly into the quantum theory.¹⁶⁹

Dirac's students could expect the Q.T. (Quantum Theory) to be true — to reflect nature — because it was as beautiful, or even more beautiful, than the classical theory. There is at this stage no mention of 'axioms' of physics, but rather of simple assumptions. These notes show that as early as 1927 Dirac was developing an aesthetics of nature centred around mathematical beauty, directly within the context of his developing quantum electrodynamics.

Three years later, on 17 November, 1930, Dirac read a paper to the Cambridge Adams Society on the 'Elements of Quantum Mechanics'. He again began with the experimental situation.

A survey of the general experimental facts about atomic phenomena shows that to obtain a theory to fit these facts is not merely a question of finding the correct laws of force between the elementary particles and then applying ordinary mechanics. There are quite general phenomena which will not fit into ordinary mechanics no matter what laws of force one assumes. For example the law of black-body radiation. [...] We are thus forced to the conclusion that the mechanics itself must be altered.¹⁷⁰

For idealists, experiment cannot force theory change. And again, Dirac presented the aesthetically pleasing character of the classical theory as an obstacle. 'The usual mechanics is so neat and simple that one is inclined to feel that it is a great pity that it has to be altered, and that any alteration one may make must be artificial and ugly'. Disturbing or changing that which is neat and simple makes it ugly and artificial. Neatness and simplicity were associated with beauty and naturalness.

Dirac would go on to further associate beauty with naturalness. He did this through a positivistic philosophy of quantum mechanics (Q.M.).

The necessity for an alteration appears in a more natural light, however, if one looks at it from the following point of view. We want our theory to be concerned with only observable things. It is usually assumed that one can see all there is to see about a mechanical system and that this will not upset the system, but is this assumption justifiable? When the things we observe are very small, like atoms, might not our observations necessarily disturb them? Q.M. is built up on the doctrine that there is a theoretical minimum to the amount of disturbance one causes, no matter how carefully one makes the observation.¹⁷¹

¹⁶⁸Ibid.

¹⁶⁹Ibid.

¹⁷⁰Dirac Papers, Series II, Box 26, Folder 4, 'Cambridge Adams Society Elements of Quantum Mechanics', 1.

¹⁷¹Ibid.

Here Dirac expressed a view akin to that expressed in Heisenberg's early writing on quantum theory, that the theory should only be concerned with the observable.¹⁷² This view was positioned as a way of naturally motivating theory change.

Pais prefers not to analyze the difference between simplicity and beauty, though Dirac himself did. Dirac most clearly expressed his views in print in his 1939 James Scott lecture.¹⁷³ In a draft typescript for this lecture he expressed his views concisely:

In the last century physicists had a mechanistic view of the universe, in which it was taken for granted that the fundamental laws of motion should be of a simple mathematical form. When Einstein put forward his theory of relativity it became necessary to change many of the laws of motion and the new ones were usually less simple. In spite [*sic*] of this physicists prefer the new laws. This is because they have greater mathematical beauty — a quality which, like beauty in art, cannot be defined, but is readily appreciated by those who have studied mathematics.

It appears therefore that mathematical beauty is really a more important requirement than simplicity in theories of physics. This is confirmed by modern developments in quantum theory, which are remarkable for the elegance of their formalism.¹⁷⁴

This was a further elaboration of his views on beauty and simplicity. While earlier naturalness, simplicity, and beauty were connected into the same conceptual sphere, here Dirac differentiated them. The James Scott lecture was also the most important public expression of his views in the 1930s.

3.2. Beauty in retrospect

Later in his life Dirac gave many historical talks on his career. In his 20 March, 1974, Loeb lecture at Harvard University 'The Road that Led to Antimatter,' Dirac set out to 'persuade' his audience 'that the formalism is really the more fundamental approach. The formalism usually came first'.

One first discovers equations and get familiar with used to them. One sees how to apply them to various physical examples. Only when one has become familiar with them does one acquire a feeling for the physical concepts that these equations involve. There is quite an interval of time — maybe even a few years — during which one has exact equations and only obscure physical concepts in one's mind.¹⁷⁵

Here Dirac used a picture of the history of his discipline (accurate or not) as a rhetorical device. He argued that because the formalism usually came first, it is more fundamental. For Dirac in these years, because it was so, it should be so.

In fragmented notes to himself for a talk at the 29 June 1976 Lindau meeting of Nobel Laureates, Dirac expressed his 'Basic beliefs and Prejudices in Physics':

My own basic beliefs To begin with Bohr orbits
They led nowhere. Heis[enberg] had different ideas
When I found I had wrong ideas I had to set up different basic ideas
Math. beauty is the underlying one
It was this that led de Broglie to his waves
The Heis[enberg] Q.M. led to a very beautiful and powerful Q.M. but it was non-rel[ativistic].
Klein-Gordon eq had $-ve$ prob[abilities].
I found a new equn with $+probs$ but $-ve$ energies
New basic beliefs about the vacuum
I had the prejudice that only 2 elem[entary] particles, so the hole had to be protons¹⁷⁶

Dirac here sketched the story of his early career. He started analyzing Bohr's orbits, taking them as physically real paths of electrons in atoms. But they led nowhere. Heisenberg's quantum theory

¹⁷²Mara Beller, 'The Rhetoric of Antirealism and the Copenhagen Spirit', *Philosophy of Science*, 63, no. 2 (Jun. 1996), 183–204.

¹⁷³Dirac, 'The Relation between Mathematics and Physics' (note 142). Published as P.A.M. Dirac, 'The Relation between Mathematics and Physics', *Proceedings of the Royal Society (Edinburgh)*, 59(Part II) (1938-39), 122–9.

¹⁷⁴Dirac Papers, Series II, Box 36, Folder 27.

¹⁷⁵Dirac Papers, Series II, Box 29, Folder 2, Harvard 'The Road that Led to Antimatter', 1.

¹⁷⁶Dirac Papers, Series II, Box 29, Folder 18, 'Lindau' 'Basic beliefs and Prejudices in Physics', 2.

pointed another way. But in the absence of physical ideas to base his work upon, Dirac needed 'different basic ideas': 'Math[ematical] beauty is the underlying one'. Dirac saw de Broglie as a kindred spirit in this regard, and he saw Heisenberg's work as beautiful, too. Dirac identified a problem with the existing Klein-Gordon equation and fixed it. This engendered new basic beliefs about the vacuum. But he saw another set of basic beliefs (that there were only two elementary particles in nature) as a prejudice that led him astray.

On 2 June 1977 at New Orleans's Loyola University, Dirac gave a talk that lucidly expresses his position, and his view of his own history.¹⁷⁷

I am glad to have this opportunity to talk to you about the math foundations of Q.T. because it allows me to express my own views on the subject. My own views differ somewhat from those of the majority of physicists. I want to emphasize the necessity for a sound math basis for any fundamental phys. theory. The maths is all important. It dominates. Whatever physical or philosophical ideas one has, one must adjust them to fit the maths. Not the other way round. Too many physicists are inclined to start with some definite phys. idea that they think should apply to fundamental physical theory

They then try to develop it to build up math equations to express it

They then run into difficulties — they find things do not fit

They may find, for example, that an integral diverges where the physics is suitable only if it converges. They ought to realize at this stage that the whole method of approach is wrong. They should have set up the maths first

The reason for this is that one is allowed to tinker with physical ideas or philosophical views

One is allowed to adjust or modify them. But one cannot do this with the maths. Maths is subject to rigid rules and it terribly restricted by strict logic

The reason I feel so strongly about these views is because I have had great success with them in the past. My early work was based on Bohr orbits and was completely unsuccessful. I was taking the Bohr orbits as physically real and trying to build up a maths for them

Now we all know the Bohr orbits are not a physically valid idea we see how futile such work is. Heisenberg opened my eyes to the need for a broader math[ematical] base.

I learned my lesson then.

Not to work from a specific physical idea, such as Bohr orbits. One should concentrate on getting an interesting maths. In this case the maths of non-commuting quantities. Just a general discussion of this maths led to the analogy between commutator and P[oisson]b[racket]s and this led to a connection between non-comm[utative] algebra and dynamics

This work led directly to a new mechanics in Q.M.¹⁷⁸

Kragh is of course correct to caution that one cannot take physicists' recollections at many decades' distance as pure statements of fact. And in the case of Dirac, Galison's work has shown that the content of his published papers is not necessarily a guide to the detailed mathematical method through which Dirac came to his results. Dirac may have used projective geometry in private, and expressed his results algebraically in public. However, despite these qualifications, this paper and Dirac's 1927 lecture notes demonstrate that Dirac's general epistemological strategy remained continuous through his most productive and also his least valued periods. Whether playing with geometry or algebra, Dirac's commitment to putting the mathematics first — and to searching for beauty there — was a constant of his scientific career. It is possible that a comparative analysis of different periods of Dirac's career could demonstrate that his later work was *more strongly* influenced by mathematical beauty. But I believe this analysis has shown that there was no 'clear dividing line' separating these periods.

Near the end of his career, and his life, Dirac was shocked by the turn of events surrounding the experimental claims to have discovered his magnetic monopole and their eventual debunking. This challenged his faith in his methods. On 28 May 1981 he wrote that 'One should conclude that pretty mathematics by itself is not an adequate reason for nature to have made use of a theory. We still need to learn in seeking for the basic principles of Nature'.¹⁷⁹

¹⁷⁷Printed in P. A. M. Dirac, 'The Mathematical Foundations of Quantum Theory', in *Mathematical Foundations of Quantum Theory*, ed. by A. R. Marlow (New York: Academic Press, 1978), pp. 1–8.

¹⁷⁸Dirac Papers, Series II, Box 29, Folder 22, 'New Orleans', '[Mathematical foundations of quantum theory]', 1–2.

¹⁷⁹Dirac Papers, Series II, Box 29, Folder 42, 'New Orleans' 'Pretty Mathematics', 2.

4. Conclusion

Dirac's epistemological strategies were part of his broader form of life. In a 1962 interview with Kuhn and Eugene Wigner, Wigner inquired after Dirac's habits at Cambridge:

Wigner: I don't think I ever had as extended a conversation with Paul as we are having now, at least not in [*sic*]. What was your daily occupation? How much did you go to lectures, how much did you sit in your room, how much did you talk to people? Did you go to theatres?

Dirac: I never went to theatres. I spent most of my time by myself, sitting working things out or going for walks. I used to spend every Sunday going for a long walk, a whole day walk, taking my lunch with me, like I did yesterday. During those long walks I would not intentionally think about my work, but I might perhaps review it. I found these occasions most profitable for new ideas coming. It was on one of those occasions that the possibility of ab-ba corresponding to a Poisson bracket occurred — on one of those Sunday walks.

Wigner: But on week days, how much time did you spend in lectures, how, much in your room?

Dirac: I don't remember just how many lectures I had. Maybe four or five a week, something of that order. I might be able to look it up . . . I have some notebooks of my lectures. But I would mainly spend the mornings and the evenings studying and took short walks in the afternoons. With a long walk all day Sundays.¹⁸⁰

Andrew Warwick has explicated the connections between long walks at Cambridge and conceptions of masculinity, that a disciplined (male) mind and a disciplined (male) body went together at Cambridge.¹⁸¹ Though Warwick's analysis ends just before Dirac came to Cambridge, we can see the same logics at work.¹⁸² Dirac can in some sense be seen as a solitary 'genius,' eschewing social contact for solitary work and walks. But this should not be taken too far. As this paper demonstrated, Dirac cultivated an elite correspondence network that informed his science. As he established himself as a researcher, he thought alone and he thought with others. Perhaps he was most comfortable on paper.

As Warwick has emphasized — and as I explored in the more-contemporary context of relativistic Penrose diagrams — theoretical physics has a social, pedagogical, and material culture.¹⁸³ In his views on approximation, explored above, Dirac emphasized the role of his engineering background in his aesthetics of physics. He found beauty in the approximate equations that engineers used to understand and manipulate the world. And this beauty became part of his metaphysics — nature was beautiful. Surely this appreciation for natural beauty and his habit of taking long walks and mountaineering were not unrelated.

Dirac's solitary work was not immaterial, the stuff of thought. It was a daily practice of inscription — of thinking while writing, drawing, and walking. In the same interview at Wigner's house, Kuhn inquired after this material record of Dirac's physics.

Kuhn: Do you have notebooks? Do you have things that go back to that period?

Dirac: I used to work on scraps of paper.

Kuhn: What happened to the scraps of paper?

¹⁸⁰Interview with P. A. M. Dirac by Thomas S. Kuhn and Eugene Paul Wigner, at Wigner's home, Princeton, New Jersey April 1, 1962, http://www.aip.org/history/ohilist/4575_1.html, [accessed 21 October 2013].

¹⁸¹Warwick, *Masters of Theory* (note 71), pp. 176–226.

¹⁸²And Dirac was not studying for the mathematical tripos exam that centres Warwick's analysis. Dirac arrived at Cambridge with a degree from Bristol as a research student studying for the PhD.

¹⁸³Aaron Sidney Wright, 'The Advantages of Bringing Infinity to a Finite Place: Penrose diagrams as objects of intuition', note 149.

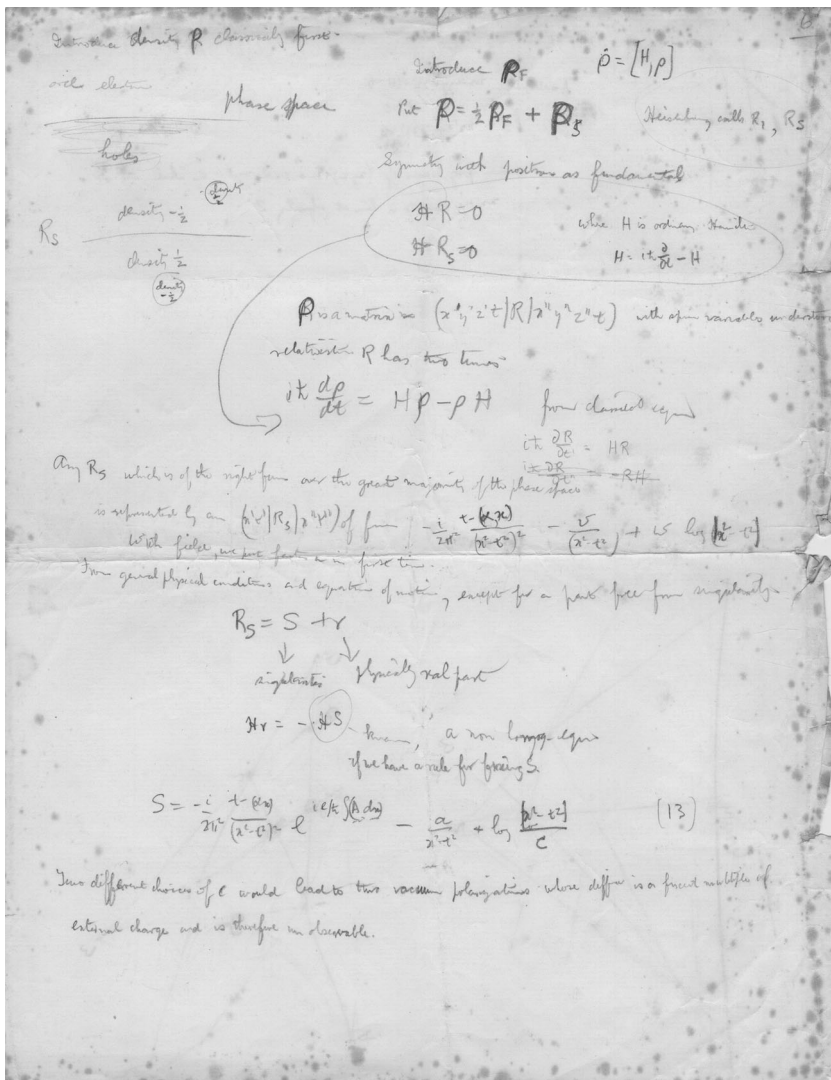


Figure 3. A page from Dirac’s archive, Dirac Papers, Series II, Box 36, Folder 22 “Theory of the Positron.” Florida State University Libraries, Tallahassee, FL.

Dirac: I’ve kept a lot of them. I’ve got some big piles of them . . . Most of it is too scrappy to be able to suggest anything. But there may be some bits which would be useful.¹⁸⁴

Dirac was correct that his habit of writing notes on scraps of paper makes much of his archive ‘too scrappy to be able to suggest anything’. Nevertheless it may be instructive to show a page from Dirac’s working life. Figures 3 and 4 are from Dirac’s work in 1934 on a theory of the positron.¹⁸⁵ This page shows Dirac drawing a picture of his infinite sea of negative-energy electrons at the top left. And it shows him setting up a Hamiltonian analysis of the density operator, ρ . On the verso, we see Dirac applying the projective geometry Galison has discussed, though Dirac’s exact procedure here is

¹⁸⁴Interview with P. A. M. Dirac by Thomas S. Kuhn and Eugene Paul Wigner, at Wigner’s home, Princeton, New Jersey April 1, 1962, http://www.aip.org/history/ohilist/4575_1.html [accessed 21 October 2013].

¹⁸⁵Dirac Papers, Series II, Box 36, Folder 22, ‘Theory of the Positron’.

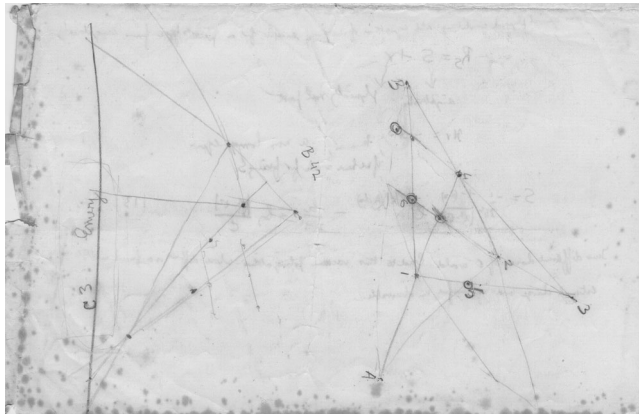


Figure 4. Detail of verso page from Dirac's archive, Dirac Papers, Series II, Box 36, Folder 22 "Theory of the Positron." Florida State University Libraries, Tallahassee, FL.

obscure.¹⁸⁶ As mentioned above, this provides evidence that Dirac continued to use his geometrical analysis beyond his training as an engineer and within his mature physics.

This sets up a hierarchy of epistemological strategies Dirac employed. Most generally, he put the mathematics first. In an untitled note from 27 November 1975 he went so far as to ascribe agency to the mathematics itself:

If you are receptive, humble the mathematics will lead you by the hand. [...]

Again and again, when I have been at a loss how to proceed, I have just had to wait until I have felt the mathematics lead me by the hand. It has led me along unexpected path, a path where new vistas open up, a path leading into new territory, where one can set up a new base of operations, from which one can survey the surroundings and plan future progress.¹⁸⁷

Dirac's vivid geographical metaphor recalls his frequent mountaineering expeditions and long country walks. At this most general level, mathematics comes first in the process of learning about the physical world. And the correct mathematics should be beautiful. We can then move to a more specific level and ask what forms of mathematical analysis Dirac used. There is good evidence that these included Hamiltonian methods of mechanics, non-commutative algebra, and projective geometry. By keeping the Hamiltonian form of mechanics at the forefront Dirac was able to transfer much of the machinery — and 'beauty' — of classical mechanics into the quantum realm. This quantum realm was inscribed through the use of non-commutative algebra.

The role of projective geometry is more difficult to discern, because Dirac did not often reflect on its role and it did not make an appearance in his published work. In his 1963 interview, Kuhn asked about a remark Dirac made that his work was largely geometrical, since his exposure to projective geometry. Dirac replied, in part:

Perhaps I didn't tell you that I kept up my connection with geometry some time after I came to Cambridge. There was a Professor Baker, a professor of geometry, who used to give tea parties on Saturday afternoons to people who were keen on geometry; after the tea someone would give a talk on some recent research work on one geometrical subject. I went to those tea-parties and absorbed quite a lot of geometry then. I talked once or twice myself. I remember I worked out a new method in projective geometry and gave a talk about that at one of these meetings. I never published this method. Well, that's a good deal about working with the geometry of four or more dimensions. Four dimensions were very popular then for the geometrists to work with. It was all done with the notions of projective geometry rather than metrical geometry. So I became very familiar with that kind of mathematics in that way. I've found it useful since then in understanding the relations which you have in Minkowski space. You

¹⁸⁶Galison, 'The Suppressed Drawing' (note 118).

¹⁸⁷Dirac Papers, Series II, Box 29 Folder 17 '[Mathematics will lead you by the hand]'.

can picture all the directions in Minkowski space as the points in a three-dimensional projective space. The relationships between vectors, null-vectors and so on — and you get at once just the relationships between points in a three-dimensional vector space. I always used these geometrical ideas for getting clear notions about relationships in relativity although I didn't refer to them in my published works.¹⁸⁸

This gives a picture of quite a lot of use of projective geometry in Dirac's career. It is important to note that by 1963 a significant portion of Dirac's career had been spent on relativity, both integrating the special theory with quantum mechanics and developing the general theory.¹⁸⁹

Dirac's daily practices cemented his epistemic strategies and his ontology. His search for beauty was rooted in daily activities of calculating, often on scraps of paper, mostly in solitude. However, his physical isolation while working was complemented by an intellectual community of correspondents with whom he thought through the repercussions of his theory. Mathematical beauty was an epistemological guide and an ontological presupposition. For Dirac, searching for beautiful mathematics worked because nature was beautiful.

Funding

The author gratefully acknowledges the support of the Social Sciences and Humanities Research Council of Canada [767-2011-2547, 756-2014-0128].

ORCID

Aaron Sidney Wright  <http://orcid.org/0000-0001-8428-890X>

¹⁸⁸Interview with Dr. P. A. M. Dirac, By Thomas S. Kuhn, At Cambridge, England, May 6, 1963, http://www.aip.org/history/ohilist/4575_2.html [accessed 27 October 2013].

¹⁸⁹P. A. M. Dirac, *General Theory of Relativity* (Princeton, NJ: Princeton University Press, [1975] 1996).