Mathematical and Causal Reasoning in Dirac's Prediction of the Positron

Sorin Bangu
Cambridge HPS
sib24@cam.ac.uk

This paper is concerned with PAM Dirac's 1928 prediction of the positron. I argue that although this prediction has been subject to many analyses by philosophers and historians of physics, a satisfactory philosophical framework able to accommodate it is still missing.[1] I first explain why Dirac's predictive reasoning is methodologically puzzling and then I outline two possible ways to approach this difficulty. More specifically, the tension I draw attention to is this. On one hand, *prediction,* as a fundamental scientific practice is typically accounted for within the deductive-nomological (DN) model[2]; moreover, it is unquestionable that this model is powerful enough to accommodate the vast majority of predictions ever made in science. On the other hand, this model can't be successfully applied to the Dirac case. After I show why this is so, I suggest that this tension can be approached in two ways. One option is to straightforwardly deny that Dirac's achievement should be recognized as a case of *prediction*, since it does not fit the standard methodology. Another, less radical, option is to dissolve the tension by allowing the supplementation of the DN model with other predictive models. However, how exactly this supplementation should be accomplished (i.e., what these other models look like) is rather unclear.

Dirac first mentions a phenomenon – amounting to the existence of a particle – "which has not been observed" in his masterful paper "The Quantum Theory of the Electron" (1928, p. 612). His reasoning proceeds along the following lines.[3] An electron with charge $-e_0$ has potential energy $-e_0\varphi$; the wave function ψ of such an electron evolves according to the Dirac equation[4]; call this equation D^e. Relatively straightforward

---

[1] Other predictions holding pride of place in the recent history of physics (such as Gell-Mann and Ne'eman's prediction of the Omega Minus particle in 1962) are subject to the same difficulty. I provide an analysis of this historical episode in my (2008).
[2] Barrett and Stanford (2006)
[3] Pais (1986, p. 348).
[4]More precisely, D^e: $ih\partial\psi/\partial t = (-ich\alpha_x\partial/\partial x + \beta m_0c^2 - e_0\varphi)\,\psi$.

manipulations of D^e lead [5] to another equation[6]; call it D^p. Once D^p is derived, the predictive argument takes this form:

(a) D^e has a physical interpretation

(b) D^e and D^p have the same mathematical form

Therefore

(c) D^p has a physical interpretation

Since the quantum theory as a whole does not preclude descriptions like D^p (Dirac 1928, p. 612-3), the wave function featuring in it might describe a new unknown entity. As D^p suggests, this particle should have the same mass as the electron (m_0) and opposite electrical charge. The experimentalist Anderson identified this particle (to be called 'the positron'; hence $D^{\mathbf{p}}$) in 1932.

After I spell out the (a) – (c) reasoning, I show that it can't be captured by the standard DN model of predictive reasoning – where the DN model construes prediction as deduction from some laws of nature and some initial conditions. I emphasize the difference between the character of prediction in the DN model (employed, for instance, in Fresnel's prediction of the bright spot or in Einstein's prediction of the bending of light rays) and the *analogical* character of Dirac's prediction. What is striking about conclusion (c) is that it just does not *deductively* follow from the premises (a) and (b); the role of these premises seems to be to circumscribe the space of possibilities, to indicate what *could* happen, not what *must* happen – unlike the typical DN cases.

So, the DN model fails to accommodate Dirac's analogical reasoning, and one might ask what is the more general significance of this failure. One possibility is to argue that it renders Dirac's result methodologically suspect. Unlike Fresnel's, Dirac's reasoning lacks deductive force, appearing to be grounded in his well-known quasi-mystical Pythagorean convictions about the power of mathematical analogies to reveal the deep structure of physical reality (see the *Intro* to Dirac 1931, p. 60). Consequently, Dirac's reasoning should perhaps be described as a lucky Pythagorean prophecy rather than an example of scientific prediction – and physics textbooks should inform their readers accordingly. Yet one could also suggest that the lesson of this failure is different: rather than suspecting Dirac's result, we should review our methodological standards (for what counts as a sound predictive reasoning). In other words, the mismatch between the DN model and Dirac's analogical reasoning should be taken to indicate that a form of methodological pluralism about prediction is more appropriate.

While I find this pluralist approach appealing, it is not clear how this rethinking should proceed. It is an open question what a predictive model able to supplement the DN one

---

[5] I detail this derivation in the appendix. See also J. Kvasnica (1966, §36-38)

[6] D^p: $ih\partial\psi^{(p)}/\partial t = (-ich\alpha_x\partial/\partial x + \beta m_0c^2 + e_0\varphi)\psi^{(p)}$.

(and thus able to accommodate Dirac's prediction) looks like. The central difficulty to find such a model is not so much that the Dirac reasoning is analogical (this is quite common in science); it is rather to provide an elucidation of how this particular kind of analogy, a *formal / mathematical* analogy, can be invested with physical significance at all. D^p is formally similar to D^e, and this at best makes D^p a *possible* description of some physical reality. The difficulty then consists in explicating what justifies the step from 'possible' to 'actual'[7] (i.e., the very predictive step), while avoiding the implausible solution offered so far in the literature, where the justification is given in terms of physicists' commitment to metaphysical principles, such as (variants of) Leibniz's Principle of Plenitude[8].

The direction I propose to follow in articulating this elucidation is to emphasize one less mysterious, and less discussed, feature of Dirac's predictive reasoning, namely his appeal to *causal* considerations.[9] This is especially important to stress since recent work on causation (e.g., in Price and Corry's 2007 collection) suggests that this notion lacks any fundamental role in modern physics. I maintain, however, that the positron story is relevant enough to cast doubts on such a view. After I isolate these considerations and identify how they intervene in his predictive argument, I argue that they were necessary (though not sufficient) to ensure its validity.

Bibliography

Anderson, C. D. (1932a) 'The apparent existence of easily deflectable positives'. *Science* 76, 238-239.

Anderson, C. D. (1932b) 'Energies of cosmic-ray particles'. *Physical Review* 41, 405-12.

Anderson, C. D., & Anderson, H. L. (1983) 'Unraveling the particle content of cosmic rays'. In L.M. Brown, & L. Hoddeson (Eds.) *The birth of particle physics* (p. 131-54). Cambridge, CUP.

Barrett, J. and Stanford, P. K. (2006) 'Prediction' in J. Pfeifer and S. Sarkar *(ed.) The Philosophy of Science: An Encyclopedia*. New York: Routledge.

Blackett, P. M. S. (1933) 'The positive electron'. *Nature* 133, 917-18.

Bueno, O. (2005) 'Dirac and the dispensability of *mathematics' Studies in History and Philosophy of Modern Physics* 36, 465-90.

[7] Norton (1995) and (2000) takes up roughly the same question when discussing some aspects of Einstein's heuristics.
[8] See for instance Kragh (1990, ch. 14).
[9] See Bueno (2005) for the first steps in this direction.

Darrigol, O. (1992) *From c-Numbers to q-Numbers*. Berkeley and Los Angeles. Univ. of California Press.

Dirac, P. A. M. (1928) ‘The quantum theory of electron’. *Proceedings of the Royal Society of London. A117*. 610-24.

Dirac, P. A. M. (1930) ‘A theory of electrons and protons’. *Proceedings of the Royal Society of London. A126.* 360-5.

Dirac, P. A. M. (1931) ‘Quantized singularities in the electromagnetic field’. *Proceedings of the Royal Society of London. A133.* 60-72.

Kragh, H. (1990) *Dirac: a scientific biography.* Cambridge, CUP

Kvasnica, J. (1966) *Physical fields*. Dover Publications.

Lovejoy, A. O. (1964) *The Great Chain of Being*. Cambridge, Mass. Harvard Univ. Press.

Norton, J. (1995) ‘Eliminative Induction as a Method of Discovery: Einstein's Discovery of General Relativity’ in J. Leplin (ed.) *The Creation of Ideas in Physics: Studies for a Methodology of Theory Construction*. Dordrecht: Kluwer, 1995, pp.29-69.

Norton, J. (2000) "'Nature in the Realization of the Simplest Conceivable Mathematical Ideas'[1]: Einstein and the Canon of Mathematical Simplicity," *Studies in the History and Philosophy of Modern Physics, 31* (2000), pp.135-170.

Pais, A. (1986) *Inward bound: Of matter and forces in the physical world*. Oxford, OUP.

Price, H., Corry, R. (2007) *Causation, Physics, and the Constitution of Reality: Russell’s Republic Revisited.* Oxford: Oxford University Press.

Schweber, S. (1994) *QED and the men who made it*. Princeton, NJ. Princeton Univ. Press.

Appendix

Consider the Dirac equation (1-dim):

$$[D^e] \qquad ih\partial\psi/\partial t = (-ich\alpha_x\partial/\partial x + \beta m_0c^2 - e_0\varphi)\,\psi$$

Taking the complex conjugate, and multiplying both sides by (–1), we obtain

$$ih\partial\psi^*/\partial t = (-ich\alpha_x\partial/\partial x - \beta m_0c^2 + e_0\varphi)\,\psi^*$$

given that that $i^* = -i$ and that α_x and β are real. (Matrices α_x and β are real as they satisfy $(p_x^2 + m_0^2c^2)^{1/2} = \alpha_x p_x + \beta m_0 c$, where p_x is the x-component of the linear momentum **p** of the particle.)

Multiplying the later relation by a matrix C such that $C\alpha_x = \alpha_x C$ and $C\beta = -\beta C$, we get

$$ih\partial(C\psi^*)/\partial t = (-ichC\alpha_x\partial/\partial x - C\beta m_0c^2 + e_0\varphi C)\ \psi^*$$

which can be rewritten as

$$ih\partial(C\psi^*)/\partial t = (-ich\alpha_x\partial/\partial x + \beta m_0c^2 + e_0\varphi)C\psi^*$$

Relabeling $\psi^{(p)} = C\psi^*$, we get

$[D^p]$ $$ih\partial\psi^{(p)}/\partial t = (-ich\alpha_x\partial/\partial x + \beta m_0c^2 + e_0\varphi)\psi^{(p)}$$