REALISM IN PHYSICS

E.B. Davies

January 2002

Abstract

We argue that the experimental verification of Newtonian mechanics and of non-relativistic quantum mechanics carry no implication that space is continuous. This provides evidence against the realist interpretation of the most mathematical parts of physics.

1 Introduction

In his influential book **The Scientific Image** van Fraassen defined scientific realism as follows.

> Science aims to give us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true, [Fr 1980].

He emphasized the inclusion of the words 'aims' and 'literally' in this definition.

A first problem with this definition is that much of science is not about the production of theories. The fact that living organisms are composed of cells is as realist a statement as one could look for in science, but its truth is a matter of observation, not of theory. Its discovery was an inevitable consequence of the development of new technology (the microscope), and was not theory driven. The same can be said about many other aspects of biology, such as the discovery of the similarities between the genetic codes of a huge range of species. In order to make sense of the definition it is best to restrict its scope to the physical sciences.

Dalton's atomic theory provides good evidence in support of scientific realism. Its truth was doubted by many of the best chemists throughout the nineteenth century, in spite of its increasing empirical success in explaining chemical reactions. Only in the twentieth century were these doubts dispelled and was the existence of atoms regarded as firmly established. Dalton's ideas about molecular structure were largely rejected by his successors, but are now regarded as unassailable fact by all chemists. The theory of molecular structure is refutable, but it has passed so many and such diverse tests that nobody has any doubt that it does provide a true insight into the world.

On the other hand, once one starts to examine fundamental physics, one's confidence in the realist picture of science starts to falter. Let us look at a more recent definition of Psillos.

> Mature and successful scientific theories are well-confirmed and approximately true of the world. So the entities posited by them, or, at any rate, entities very similar to those posited, inhabit the world, [Ps 2000].

There are three obvious candidates for such theories: Newtonian mechanics, quantum theory and general relativity. The first has supposedly been superseded, but is still as heavily used as before this happened. The other two both reign supreme within their domains of applicability, but they have no consistent amalgamation. Each involves an entirely different understanding of the nature of space (or space-time) and a very different idea about what constitutes a material body or particle. They depend upon entirely different branches of mathematics. Only quantum theory can even begin to explain the interference effects observed in the double slit experiment.

The objective existence of atoms and molecules does not imply that their current description using the mathematics of quantum theory is the most fundamental one. The solitary water waves first observed by John Scott Russell are now explained as soliton solutions of certain non-linear ordinary differential equations. However, there is a quite different explanation of them as collective excitations of vast numbers of water molecules. Similarly electrons and atoms may be quasi-particles: collective excitations of some quite different entity. The current ferment in fundamental physics only guarantees that they are not what they currently seems to be. Logically speaking, this does not contradict van Fraasen's definition of scientific realism, but the aim may be at such a distant target that we can never know if it has been hit.

The belief of many physicists that the final goal of science is to construct a completely objective 'Theory of Everything' has been challenged in a deliberately controversial paper by Laughlin and Pines, the former a Nobel Laureate in physics. They argue that attempts to derive what we know from a Theory of Everything could not succeed because the relevant calculations could not be performed even if the theory existed. They also express doubts that current progress towards such a theory has the significance claimed.

> Thus the existence of (the Josephson quantum and the quantum Hall effect) is profoundly important, for it shows us that for at least some fundamental things in nature the Theory of Everything is irrelevant. P. W. Anderson's famous and apt description of this state of affairs is "more is different". The emergent physical phenomena regulated by higher organizing principles have a property, namely their insensitivity

to microscopics, that is directly relevant to what is knowable in the deepest sense of the term, [LP 2000].

Our goal in this paper is to show that one can produce mathematical models whose predictions are the same as those of more standard theories, but in which the structure of space is quite different. We show that quantum mechanics provides no reason to believe that physical space is continuous, even though this belief underlies most current research in physics. Moreover the apparent rotational symmetry (isotropy) of space does not imply any underlying symmetries of a quantum mechanical model which describes motion within it; in the current jargon one should say that the isotropy is an emergent phenomenon. These two facts suggest that it is better to adopt an empirical attitude towards highly mathematical theories of the world: their predictive success does not carry the expected implications about what 'the world is really like'.

We proceed by constructing an alternative model of non-relativistic quantum mechanics, in which physical space does not have the usual Euclidean structure. The numerical predictions of this new model are indistinguishable from those of the standard model, because it is based upon a method used for numerical calculations. We are not claiming that the new model has any special merits: our point is that its *logical and philosophical* relationship with the Euclidean picture is quite different from that of the standard model, and in many situations there is no good reason to prefer one model to the other. The constructions involved are completely standard mathematically, and only the interpretation has any degree of novelty. A feature of our model which distinguishes it from the standard version of quantum theory is that it cannot be constructed by 'quantization' of a preceding classical model – there is no classical dynamics on a discrete configuration space. There is a classical limit, but this takes place in a space different from the physical space of the model.

Minsky has shown that some aspects of the classical theory of fields can be simulated on a discrete space-time by cellular automata which do not recognize the real number system,[Mi 1982]. The idea that space-time might ultimately be discrete or might have a very small-scale structure quite unlike that of Minkowski space is well-known to those working in fundamental physics. One purpose of this article is to present an elementary version of such ideas so that their significance can be explored without getting lost in technicalities. The results described in this paper support the philosophical position called entity realism, discussed briefly in the final section.

2 The Model

Both Newtonian mechanics and non-relativistic quantum mechanics assume that physical space is represented mathematically by the three-dimensional Euclidean space, $\mathbf{R}^3$. In this section we replace this by $S = \mathbf{Z}^3$, the space of triples of integers.

Since physical space in this model is discrete there can be no direct analogue of Newtonian mechanics in it: the only continuous paths are fixed.

It is normal to regard S as a graph in which two points $m = (m_1, m_2, m_3)$ and $n = (n_1, n_2, n_3)$ are taken to be neighbours if the graph distance

$$d_1(m, n) = |m_1 - n_1| + |m_2 - n_2| + |m_3 - n_3| \tag{1}$$

equals 1. This distance does *not* approximate to the Euclidean distance

$$d_2(m, n) = \sqrt{(m_1 - n_1)^2 + (m_2 - n_2)^2 + (m_3 - n_3)^2}$$

as the points m, n separate, so if we let $\delta > 0$ denote the edge length of the graph, it is not obvious in what sense δS converges to Euclidean space as $\delta \to 0$. This question is addressed in Section 4. This is based upon [Da 1993], which constructs a new distance d_3 on S which is both intrinsic and asymptotically Euclidean. It is an interesting fact that several aspects of the analysis of graphs seem to depend upon the use of Connes' non-commutative geometry, indicating that discretization is not a trivial matter, [Da 1993]. Indeed the graph-theoretic analogue of a problem on a manifold is often harder than the original problem.

In this paper we regard δ as a physical constant, describing the distance scale at which physical space becomes discrete. This scale must be much smaller than the diameter of an atom in order to guarantee consistency with what has already been verified using the standard model of quantum theory.

The only symmetries of our space S are translations by integer amounts, rotations through integer multiples of $\pi/2$ and certain reflections. In particular S does not possess anything remotely like the full symmetry group of Euclidean space, however small $\delta > 0$ is chosen to be. Indeed, it would be possible to modify the definitions in this section, replacing S by a countable, irregular array of points in $\mathbf{R}^3$ with constant average density, subject to suitable conditions. If one did this there would be no non-trivial symmetries of S.

Although classical dynamics cannot exist in this context, it is easy to produce a version of quantum mechanics. We do this by adapting the systematic approach of Mackey to the discrete context, [Ma 1963]. For a single free particle the Hilbert space is taken to be $l^2(\mathbf{Z}^3)$ and the Schrödinger equation is

$$i\frac{\mathrm{d}f}{dt} = Hf = -\frac{c}{2}\Delta f$$

where Δ is the usual three-dimensional discrete (graph) Laplacian, given by

$$\Delta f(m) = \sum_{r=1}^{3} \{f(m + e_r) + f(m - e_r) - 2f(m)\},$$

$\{e_r\}_{r=1}^3$ being the standard basis of $\mathbf{Z}^3$. The normalization constant c is determined by a standard procedure, and we choose physical units so that $c = 1$. The time evolution is given by

$$f(t) = \mathrm{e}^{-iHt} f(0)$$

and defines a one-parameter unitary group on $l^2(\mathbf{Z}^3)$.

The extension of the above definitions to several particles obeying Bose or Fermi statistics and interacting by means of typical two-body potentials is straightforward. The existence of atoms holds in the model exactly as it does in the standard model, and their energy levels are almost exactly the same in the two cases. Indeed what we are describing is a standard discretization of quantum theory which might, in principle, be used for numerical computations. It would be highly inefficient for this purpose, but there is no issue of principle. Our viewpoint is different, since we take a single value of $\delta > 0$, which is supposed to be small enough for its effects not to be detectable at the present time.

Although there can be no continuous motion in $\mathbf{Z}^3$, there is a different way of introducing classical dynamics related to the above quantum evolution. For each $r = 1, 2, 3$ there are 'position' and 'momentum' operators, defined by

$$\begin{aligned} Q_r f(m) &= m_r f(m) \\ P_r f(m) &= \frac{f(m+e_r) - f(m-e_r)}{2i}. \end{aligned}$$

A routine calculation in the Heisenberg picture shows that

$$\begin{aligned} i[H, Q_r] &= P_r \\ i[H, P_r] &= 0 \end{aligned}$$

for all r. Hence if the expected position and momentum of a Schrödinger picture quantum state f with $\|f\|_2 = 1$ are defined by

$$\begin{aligned} q_r(t) &= \langle Q_r f(t), f(t) \rangle \\ p_r(t) &= \langle P_r f(t), f(t) \rangle \end{aligned}$$

we see that the momentum is independent of time, while

$$q_r(t) = q_r(0) + t p_r(0)$$

for all $t \in \mathbf{R}$ and $r = 1, 2, 3$. This is Newton's first law of motion.

If $\mathcal{F} : l^2(Z^3) \to L^2((-\pi, \pi)^3)$ is the usual unitary Fourier transform operator, and $\theta = (\theta_1, \theta_2, \theta_3) \in (-\pi, \pi)^3$, then

$$\mathcal{F} P_r f(\theta) = \sin(\theta_r) \mathcal{F} f(\theta).$$

It follows that $-1 \le p_r \le 1$ for any normalized state $f \in l^2(\mathbf{Z}^3)$. In other words there is a maximum speed of propagation in this model. However, if the value of δ is sufficiently small, this speed will be so great that it will be unattainable in any imaginable circumstances. At low energies (i.e. for small θ) we have

$$\mathcal{F} P_r f(\theta) \sim \theta_r \mathcal{F} f(\theta)$$

and one recovers the standard commutation relations for $L^2(\mathbf{R}^3)$.

The kinetic energy operator H is given in momentum space by

$$\begin{aligned} Hf(\theta) &= \frac{1}{2}\sum_{r=1}^{3}\{2-2\cos(\theta_r)\}\, f(\theta) \\ &= h(\theta)f(\theta) \end{aligned} \tag{2}$$

The function $h(\cdot)$ does not have rotational symmetry. However at low energies we have

$$h(\theta) \sim \frac{1}{2}\left(\theta_1^2+\theta_2^2+\theta_3^2\right).$$

The isotropy in this limit is seen to be an elementary consequence of the fact that any smooth function is approximately quadratic near a local minimum. It is an emergent phenomenon of great importance for computations, but with no deep significance.

3 Conservation of Momentum

The law of conservation of momentum for interacting particles is less straightforward than might be imagined. For simplicity we consider only the case of two one-dimensional particles with unit mass, so that the relevant Hilbert space is $l^2(\mathbf{Z}\times\mathbf{Z})$; the ideas can easily be extended to more particles and higher dimensions. We suppose that the Hamiltonian is $H = H_0 + V$, where the kinetic energy operator H_0 commutes with all space translation operators

$$T_a f(m,n) = f(m+a, n+a)$$

and V is a two-body potential:

$$Vf(m,n) = V(m-n)f(m,n).$$

It is immediate that

$$[T_a, V] = [T_a, H_0] = 0$$

for all $a \in \mathbf{Z}$, and hence that the total momentum operator

$$P = \frac{1}{2i}\left(T_1 - T_{-1}\right)$$

commutes with H. In other words the total momentum is a conserved quantity. On the other hand P is not equal to $\tilde{P} = P_1 + P_2$, where P_r are the momenta of the individual particles, and $\tilde{P}$ is not a conserved quantity. So the momentum of a group of particles is not additive. However the expected values of P and $\tilde{P}$ are almost equal for any state which varies slowly in space (by comparison with the size of δ).

The centre of mass of the system corresponds to the operator

$$Q = \frac{1}{2}\left(Q_1 + Q_2\right)$$

and it would be nice to have

$$i[H,Q]=\frac{1}{2}P$$

or equivalently

$$i[H_0,Q]=\frac{1}{2}P \qquad (3)$$

since this would immediately imply that the expected value of the centre of mass moves with constant speed in any state. However this is false even for semi-relativistic classical Hamiltonian functions such as

$$H(p_1,p_2,q_1,q_2)=\sqrt{m^2+p_1^2}+\sqrt{m^2+p_2^2}+V(q_1-q_2)$$

so it is optimistic to expect it in quantum mechanical situations unless the kinetic energy is a quadratic function of the momentum, which in our case it is not.

So far we have not specified H_0 apart from requiring that it commutes with translations T_a. If we make the obvious definition

$$\begin{aligned} H_0 f(m,n) &= -\frac{1}{2}\{f(m+1,n)+f(m-1,n) \\ &\quad + f(m,n+1)+f(m,n-1)-4f(m,n)\} \end{aligned}$$

then (3) is not satisfied. On the other hand for

$$\begin{aligned} H_0' f(m,n) &= -\frac{1}{4}\{f(m+1,n+1)+f(m-1,n+1) \\ &\quad + f(m+1,n-1)+f(m-1,n-1)-4f(m,n)\} \end{aligned}$$

one may easily verify (3). Both operators commute with translations and are discretizations of the usual continuous kinetic energy operator. So it appears that one has to make a choice between additivity of the kinetic energy and uniform motion of the centre of mass. The fact that different discretizations of the Laplacian may have significantly different degrees of isotropy is well known, and is explored in the fluid mechanics context in [RB 2001, p 26, 128, 205]. For any state which varies slowly as a function of space, the two operators are practically indistinguishable.

We finally outline a possible approach to the derivation of Newtonian mechanics from our model. If one follows the method of Landau, [La1 1994, sect. 2.4], [La2 1996], then we conjecture that one would obtain classical dynamics with respect to the Hamiltonian function

$$H(p_1,p_2,q_1,q_2)=h(p_1)+h(p_2)+V(q_1-q_2) \qquad (4)$$

where $h(\cdot)$ is defined in (2). Newtonian mechanics would then be obtained by taking a second low momentum limit. Only at this stage would the full Euclidean invariance emerge.

4 The Asymptotic Isometry Group

Our physical space S has a very limited isometry group. The scaling limit as $\delta \to 0$ of $\delta \mathbf{Z}^3$ with the metric δd_1 is $\mathbf{R}^3$ with the metric d_1. Here we are taking limits of metric spaces in the sense of Gromov-Hausdorff, [BH 1999, p 72]. The limit space has full translational symmetry, but its rotational group is finite of order $3! \times 2^3$. This type of limit is relevant to the macroscopic structure of substances such as diamond. Their atomic structure cannot be seen macroscopically, but still affects their gross structure through preferential cleavage planes.

Nevertheless the full isometry group of Euclidean space may be recovered from S by using a different approach. Let X be any set with a metric d, and define an approximate contraction g of X to be a map $g : X \to X$ for which there exists a constant $c = c_g < \infty$ such that

$$d(gx, gy) \leq d(x, y) + c$$

for all $x, y \in X$. The set T of all approximate contractions is clearly a semigroup with an identity element e.

Now define the equivalence relation $\sim$ on T by $g \sim h$ if there exists a constant $c = c_{g,h} < \infty$ such that

$$d(gx, hx) \leq c$$

for all $x \in X$. It is elementary to verify that $g_1 \sim g_2$ and $h_1 \sim h_2$ imply $g_1 h_1 \sim g_2 h_2$. Finally define G to be the set of all $g \in T$ such that there exists $h \in T$ with $gh \sim hg \sim e$. Clearly G is a sub-semigroup of T which possesses inverses up to $\sim$. If $\tilde{G}$ is the space of equivalence classes of G then $\tilde{G}$ is a group with respect to the inherited multiplication, which we call the asymptotic isometry group of X with respect to the metric d. Geometrically speaking $\tilde{G}$ is obtained by ignoring finite discrepancies from the exact definition of isometry.

In the particular case $S = \mathbf{Z}^3$, we do not take d to be the graph distance d_1 defined in (1). We commented that the intrinsic metric d_3 defined in [Da 1993] is asymptotically equal to the Euclidean distance, and take d to be the Euclidean distance on $\mathbf{Z}^3$. (Theorem 6.4 of [Da 1993] is only stated in the two-dimensional case, but the method has more general applicability.) We also define $p : \mathbf{R}^3 \to \mathbf{Z}^3$ to be the projection which takes any point $x \in \mathbf{R}^3$ to the point of $\mathbf{Z}^3$ closest to x; if there is more than one such point then px is chosen from the possibilities arbitrarily. For any Euclidean isometry $u : \mathbf{R}^3 \to \mathbf{R}^3$ we may now define the associated $g \in \tilde{G}$ to be the equivalence class of the map $m \to pum$ from $\mathbf{Z}^3$ to $\mathbf{Z}^3$. The map itself need be neither one-one nor onto, but it lies in T by an elementary calculation. The map $u \to g$ from the group of Euclidean isometries of $\mathbf{R}^3$ to the group of asymptotic isometries of $\mathbf{Z}^3$ is a group homomorphism, and we conjecture that it is bijective.

We emphasize that the group $\tilde{G}$ above is a constructed quantity which does not act on S. We have shown that $\delta \mathbf{Z}^3$ with the Euclidean metric d_3 converges to Euclidean space as $\delta \to 0$; in this sense our choice of physical space is similar to the standard one as Psillos requires. Note, however, that this is only true if one uses the Euclidean metric or the intrinsic metric d_3 on $\mathbf{Z}^3$. It is not true for the more obvious graph metric d_1. It appears from (4) that which limit is more appropriate depends on whether one is taking a zero energy limit at the same time as the continuum limit.

5 Conclusions

The above model does not avoid the use of continuous analysis, but this enters via the assumption that the phases of wave functions are complex numbers. This is logically quite distinct from the assumption that physical space is a continuum. It provides a different resolution of Zeno's arrow paradox from the usual one: in our model particles move continuously by utilising quantum superpositions between discretely separated positions. By evaluating the *expected values* of the positions of particles we found that these are continuous and obey Newton's first law, even though the *actual* positions are discrete. We also proved the equivalent of Newton's third law, namely the conservation of the *expected value* of the momentum for interacting particles on the derived, continuous version of space.

The existence of this model shows that the usual continuous notion of space may exist only by construction. In our model the full Euclidean symmetry group is an emergent property in the low momentum and large distance limit and is not a property of our physical space itself. In other words the fact that we *appear to see* certain symmetries in the world around us does not carry any implication that there are such symmetries in the world itself.

Our model is not a serious candidate as a fundamental theory of the world. Like the standard version of quantum theory it does not incorporate any ideas from the theory of relativity. A more fundamental theory need not be relativistically invariant and might be built upon a discrete version of space-time. All that is required is that some version of relativity theory should emerge from it is a suitable limit, in the same way as the Euclidean symmetry group emerges from our model.

Even if there were experimental evidence in favour of our model, this would not imply the eventual demise of the standard model. The latter has great advantages computationally: the hydrogen atom and harmonic oscillator are unlikely to be exactly soluble in our model, because of its lack of Euclidean symmetries. But this would be a matter of convenience rather than one of principle. We quote Roger Newton.

> Conceivably, the resulting predictions could be made without using the concepts and results of the theory of groups; at the very least, however, to arrive at them in such a way would take enormously more time and effort. It is not that Nature itself makes use of group theory; it is that we humans need this invaluable mental crutch to understand Nature. Mathematics is not embedded in the structure of reality, but we require

the help of its power to penetrate and describe that reality, [Ne 1997, p 141].

An overwhelming body of evidence supports scientific realism in chemistry, but our model indicates that it is much harder to do so in fundamental physics. The philosophical position which accords most closely to our results is called entity realism, invented by Cartwright and Hacking (independently) in 1983. This is currently a minority position, and allows the objective existence of unobservable entities such as electrons and atoms. It also accepts phenomenological laws about them, but adopts an antirealist stance towards fundamental laws, [Cl 2001]. However, one should not underestimate the ingenuity and creativity of theoretical physicists. Within the next twenty years they may well come up with a consistent way of unifying the known forces of nature. Whether this will be the Theory of Everything will be quite another matter.

Acknowledgments I would like to thank L J Landau and R F Streater for valuable discussions.

References

[BH 1999] M R Bridson, H Haefliger: Metric Spaces of Non-Positive Curvature. Springer, Berlin.

[Cl 2001] S Clarke: Defensible territory for entity realism. Brit. J. Phil. Sci. 52, 701-722.

[Da 1993] E B Davies: Analysis on graphs and noncommutative geometry. J. Funct. Anal. 111, 398-430.

[Fr 1980] B C van Fraassen: The Scientific Image. Clarendon Press, Oxford.

[La1 1994] L J Landau: Observation of quantum particles on a large space-time scale. J. Stat. Phys. 77, 259-310.

[La2 1996] L J Landau: Macroscopic observation of a quantum particle in a slowly varying potential. Ann. Phys. 246, 190-227.

[LP 2000] R B Laughlin and D Pines: The theory of everything. PNAS 97, 28-31.

[Ma 1963] G W Mackey: Mathematical Foundations of Quantum Mechanics. Benjamin, USA.

[Mi 1982] M Minsky: Cellular vacuum. Intern. J. Theor. Phys. 21, 537-551.

[Ne 1997] R Newton: The Truth of Science. Harvard Univ. Press, Cambridge, MA.

[Ps 2000] S Psillos: The present state of the scientific realism debate. Brit. J. Phil. Sci. 51, 705-728.

[RB 2001] J-P Rivet, J P Boon: Lattice Gas Hydrodynamics. Camb. Univ. Press, Cambridge, England.

Department of Mathematics
King's College
Strand
London WC2R 2LS
England

e-mail: E.Brian.Davies@kcl.ac.uk