

The computational status of physics

A computable formulation of quantum theory

Mike Stannett

Published online: 28 February 2009
© Springer Science+Business Media B.V. 2009

Abstract According to the Church-Turing Thesis (CTT), effective formal behaviours can be simulated by Turing machines; this has naturally led to speculation that *physical* systems can also be simulated computationally. But is this wider claim true, or do behaviours exist which are strictly *hypercomputational*? Several idealised computational models are known which suggest the possibility of hypercomputation, some Newtonian, some based on cosmology, some on quantum theory. While these models' physicality is debatable, they nonetheless throw into question the validity of extending CTT to include *all* physical systems. We consider the physicality of hypercomputational behaviour from first principles, by showing that quantum theory can be reformulated in a way that explains why physical behaviours can be regarded as 'computing something' in the *standard computational* state-machine sense. While this does not rule out the physicality of hypercomputation, it strongly limits the forms it can take. Our model also has physical consequences; in particular, the *continuity of motion* and *arrow of time* become theorems within the basic model.

Keywords Hypercomputation · Quantum theory · Theory of computation · Philosophy of science · Arrow of time · Discrete time · Natural computation

1 Introduction

According to the Church-Turing Thesis (CTT), all effective computational behaviours can be simulated by Turing machines (Kleene 1952). Although CTT was proposed in the context of formal mathematical systems, it is widely accepted that it can be applied more generally; in particular, given that physical devices are routinely used for computational purposes, it is now widely assumed that all (finitely resourced, finitely specified) physical machine behaviours can be simulated by Turing machines. However, this extended

M. Stannett (✉)
Department of Computer Science, Regent Court, 211 Portobello Street, S1 4DP Sheffield, UK
e-mail: M.Stannett@dcs.shef.ac.uk

claim¹ (known in the philosophy and computer science literature as *Thesis M* (Gandy 1980; Copeland 2002), and in physics literature as the *physical Church-Turing Thesis*; see e.g. (Deutsch 1985; Penrose 1990) and references therein) is not by any means a logical consequence of CTT, since it is not clear that every physical machine can meaningfully be said to ‘compute something’ in the same sense as Turing machines. Proponents of *digital physics* (Wolfram 2002; Lloyd 2006; Tegmark 2008) stretch CTT still further, interpreting it to mean that *all* physical behaviours (whether machine-generated or not) are Turing-simulable.

The main aim of this paper is to investigate Thesis M and its extensions in more detail. Is it actually true that all physical behaviours are necessarily computable, or are there behaviours which go beyond the Turing limit? We will show that quantum theory can be reformulated in a way that partially resolves this question, by explaining why physical behaviours can indeed *always* be regarded as ‘computing something’ in the strict state-machine sense. While our approach does not rule out the possibility of hypercomputation completely, it limits the form such hypercomputation must take.

As we recall in Sect. 2, this question has been debated indirectly over many decades (Stannett 2006); but it has become prominent recently with the rise of quantum computation and digital physics. As is well known, Shor’s (1997) algorithm can factorise integers faster than any Turing program, and this already suggests that quantum theory has super-Turing potential. However, we need to distinguish carefully what we mean by ‘hypercomputation’ in this context. Where a computational model—for example, Deutsch’s (1985) Universal Quantum Computer (UQC)—computes the same class of functions as the Turing machine, albeit potentially faster, we call it a *super-Turing* model. If it is capable of computing functions which *no* Turing machine can compute, we call it *hypercomputational*. In particular, then, while the UQC is an apparently super-Turing model, it is well known that it is not hypercomputational, whence its implementation would not resolve the question whether hypercomputation is physically feasible.

1.1 Layout of the paper

We begin in Sect. 2 by considering briefly what is already known concerning the relationship between physics and (hyper)computation. After summarising the information-theoretic approach familiar from *It from Bit*, we review three known hypercomputational systems: non-collision singularities in the Newtonian n -body problem; the Swansea *Scatter Machine Experiment* (also Newtonian); and Hogarth’s cosmologically inspired family of *SAD* computers. We then focus on quantum theory, where it is unclear whether any hypercomputational model has yet been established. The question then arises whether a new approach might be able to resolve the issue. We will show that this is indeed the case, though only to a limited extent, by deriving a first-principles reformulation of Feynman’s path-integral model; we review the standard formulation briefly in Sect. 3, and present our new formulation in Sect. 4.

In our version of Feynman’s model, there is no such thing as a continuous trajectory. Instead, whenever a particle moves from one spacetime event to another, it does so by performing a finite sequence of ‘hops’, where each hop takes the particle directly from one location to another, with no intervening motion. Although this seems somewhat iconoclastic, we argue that ‘finitary’ motion of this kind is the only form of motion actually supported by observational evidence.

¹ Andr eka et al. (2008) argue that the physical variant of CTT was first considered as far back as the 1930s.

In Sect. 5 we consider the computational significance of the model, insofar as it addresses the question whether hypercomputation is physically feasible. From a mathematical point of view it makes little difference whether we allow ‘hops’ to move a particle backwards as well as forwards in time, and we consider both models. In each case, the motion of a particle from one location to another generates a finite state machine (technically, an extended form of FSM called an *X-machine* (Eilenberg 1974)), where the machine’s states are spacetime locations, and its transition labels reflect the (classical) action associated with each hop. In unidirectional time, the regular language generated by such a machine comprises just a single word, but if we allow time to be bidirectional, the availability of loops ensures that infinite regular languages can be generated. In both cases, when the motion is interpreted as an *X-machine*, the function computed by the motion can be interpreted as an amplitude, and if we sum the amplitudes of all machines with a given initial and final state, we obtain the standard quantum mechanical amplitude for the particle to move from the initial to the final location.

Section 6 concludes our argument, and includes suggestions for further research. We note in particular that certain assumptions inherent in Feynman’s original model must be regarded as *provable theorems* of the model presented here; this includes both the *continuity of observed motion* and the *arrow of subjective time*.

2 Motivation

In this section we review various arguments both for and against the physical feasibility of hypercomputation, and its converse, digital physics; for a more complete discussion of hypercomputational models, readers are invited to consult our earlier surveys of the field (Stannett 2003, 2006). The question, whether hypercomputational behaviours are physically feasible, obviously depends on one’s conception of physics itself. Hypercomputational systems have been identified with respect to both relativistic and Newtonian physics. Where quantum theory is concerned, however, the situation is less clear cut.

2.1 Digital physics

Proponents of digital physics argue that the Universe *as a whole* is essentially computational, in the sense that its entire history can be viewed as the output of a digital computation (Schmidhuber 1997). The underlying idea appears first to have been proposed by Zuse, who suggested as early as 1967 that the Universe might be computed by a deterministic cellular automaton inhabited by ‘digital particles’ (Zuse 1967, 1969).

Wheeler’s subsequent (1990) *It from Bit* conception reflected his conviction that information is just as physical as mass and energy, and indeed the relationship between information and gravitation has remained central to theories of quantum gravity ever since Bekenstein (1972, 1973) realised that black holes must possess intrinsic entropy. Likewise, Hawking’s observation that black holes can evaporate (Hawking 1974) forces us to ask what happens to quantum correlations that previously existed between particles on either side of the event horizon? Quantum theory appears to be inconsistent with causality in such a situation (Susskind and Lindesay 2005).²

² There is as yet no empirical evidence that Hawking radiation, the mechanism by which evaporation takes place, exists in Nature. However, the final stages of a primordial micro black hole’s evaporation should

The *It from Bit* doctrine focusses on the relationship between observation and information. Just as observations provide information, so information can affect observations, as was graphically illustrated (at first theoretically and eventually experimentally) by Wheeler's famous 'delayed-choice experiment', a modified version of the dual-slit experiment. As is well known, if one slit in a barrier is covered over, photons passing through the apparatus behave like particles, but when both slits are opened the 'particles' demonstrate interference effects. Wheeler asked what would happen if the decision to cover or uncover a slit were made *after* the photon had passed through the barrier, but before the outcome were detected. In practice, the photon's behaviour reflects the decision the experimenter will eventually make, even though this decision occurs after the encounter with the barrier has taken place. This suggests that the outcome of an experiment involves an interaction between the apparatus and the observer; the results you get are in some sense changed by the questions you decide to ask; or as Wheeler put it, "Every 'it'—every particle, every field of force, even the spacetime continuum itself—derives its function, its meaning, its very existence entirely—even if in some contexts indirectly—from the apparatus-elicited answers to yes-or-no questions, binary choices, bits" (Horgan 1991).

Schmidhuber (1997, 2000) has investigated a model of physics in which all possible realities are the outcomes of computations. By considering algorithmic complexity, we can examine the probability that a randomly selected universe would conform to any given set of behaviours; specific physical situations can be examined and predictions made, some of which might, in principle, be subject to experimental verification. It is important to note, however, that the type of physics this model generates is *not* generally consistent with conventional wisdom. For example, because digital physics assumes that universes are inherently deterministic, Schmidhuber's model rejects the notion that beta decay is truly random. Similarly, his model suggests that experiments carried out on widely separated, but initially entangled, particles, should display non-local algorithmic regularities, a prediction which, he notes, 'runs against current mainstream trends in physics'.

A related concept is Tegmark's *Mathematical Universe Hypothesis*. Tegmark (2008) notes that, if a complete Theory of Everything (TOE) exists, then the Universe must necessarily be a mathematical structure. In essence, this is because a *complete* TOE should make sense to any observer, human or otherwise, whence it ought to be a formal theory devoid of 'human baggage'; consequently the TOE (and hence the Universe it specifies) is a purely mathematical structure. While this argument can obviously be challenged—it is entirely possible that pure mathematics is itself a form of human baggage and that the concept 'mathematical structure' has no meaning to creatures whose brains have evolved differently to our own—Tegmark shows that it entails a surprisingly wide range of consequences, but interestingly, these do *not* include computability. Rather, Tegmark introduces an additional *Computable Universe Hypothesis*, according to which the relations describing the Universal structure can be implemented as halting computations. This is similar to Schmidhuber's model, except that it is the relationships between objects that are deemed computable, rather than their evolution through time.

Footnote 2 continued

theoretically result in a burst of gamma-rays; one of the goals of the GLAST satellite, launched by NASA on 11th June 2008, is to search for such flashes.

2.2 Examples of physical hypercomputation

A key feature of the digital physics models described above—as well as, e.g. Zizzi's (2004) loop quantum gravity model—is that the models take the assumption of an information- or computation-based universe as their *starting point*, and then ask what consequences follow. This is inevitable, since the authors are ultimately interested in identifying experiments which might provide evidence in support of (or which falsify) their models. Clearly, however, if experiments are to distinguish between digital physics and 'conventional wisdom', it must first be necessary that digital physics and the standard model are not equivalent. It follows, therefore, that digital physics cannot tell us about the feasibility or otherwise of hypercomputation in 'standard' quantum theory.

Unfortunately, this is precisely the question we wish to answer. Rather than invent a *new* model of physics that is computational by fiat, we wish to determine whether the *standard* model is computational. Our approach, which we outline in some detail in Sects. 3 and 4, is to reformulate (a small part of) the existing model in such a way that its computational nature becomes intuitively obvious. Before doing so, however, we should explain why this task is worth undertaking—as Zuse (1969) put it, "Is Nature digital, analog or hybrid? And is there essentially any justification for asking such a question?"

2.2.1 Newtonian models (and a challenge to digital physics)

It is not often appreciated that standard Newtonian physics supports both super-Turing and hypercomputational behaviours, but as Xia (1992) has shown, the Newtonian n -body problem exhibits 'non-collision singularities', solutions in which massive objects can be propelled to infinity in finite time. This is particularly problematic for those models of digital physics which claim the Universe is generated by essentially local interactions, like those connecting processes in a cellular automaton, because the laws of physics are typically considered to be time-reversible. Consequently, if a particle can be propelled *to* infinity in finite time, it should also be possible for a particle to arrive *from* infinity in finite time. Clearly, however, there is no earliest time at which such an emerging particle first arrives in the Universe (the set of times at which the emerging particle exists does not contain its greatest lower bound). Consequently, if all objects in the Universe have finite extent and finite history, the particle's 'emergence at infinity' must involve some non-local form of interaction between infinitely many of these objects. On the other hand, Xia's model depends implicitly on an idealised version of Newtonian physics, in which gravitationally bound point-masses can approach arbitrarily closely (some such idealisation is unavoidable, as the system needs to supply unbounded kinetic energy to the escaping object as it accelerates away to infinity). While this means that Xia's result doesn't actually undermine the case for digital physics in 'real-world' terms, it reminds us that the situation is considerably more complicated than it might at first appear.

A recent series of investigations, reported in Beggs et al. (2008), concerns a collision-based computational system called the *Scatter Machine Experiment* (SME), in which a projectile is fired from a cannon at an inelastic wedge in such a way that it bounces into a detector either to one side (*up*) of the apparatus or the other (*down*); if the projectile hits the vertex, various scenarios can be posited. The wedge is fixed in position with its vertex at some height x whose binary expansion we wish to compute. The cannon can also be moved up and down, but whereas x can take any real value, we only allow the cannon to be placed at heights u which can be expressed in the form $u = m/2^n$ for suitable m and n . By repeatedly firing and then re-aligning the cannon, we can attempt to compute the binary

expansion of x , one digit at a time. The class of sets which are decidable in polynomial time, when a certain protocol is used to run the SME, is exactly $P/poly$ (the complexity class of languages recognised by a polynomial-time Turing machine with a polynomial-bounded advice function). Since $P/poly$ is known to contain recursively undecidable languages (Goldreich 2008), it follows that the scatter machine experiment—despite its evident simplicity—is behaving in a hypercomputational way.

2.2.2 Relativistic models

The SAD_n hierarchy is a family of computational models which exploit the properties of certain singularities in *Malament–Hogarth* spacetimes (Hogarth 1992; Etesi and Némethi 2002). These are singularities with computationally useful properties; in particular, if a test particle falls into the singularity, it experiences infinite proper time during its journey; but an outside observer sees the entire descent occurring in finite time. By exploiting such a singularity, we can easily solve the Halting Problem. For suppose we want to know whether some program P halts. We set it running on a computer, and then send that computer into the singularity. From our vantage point, the entire process lasts just a finite length of time, say T seconds. From the computer's point of view the descent takes forever, so if P is going to halt, it will have enough time to do so. We therefore program the computer's operating system so that, if P halts, a rocket is launched—this is possible for this kind of singularity—so as to arrive at some previously determined place and time, somewhat more than T seconds (from our point of view) after the computer is launched. We then travel to the rendezvous point. If a rocket arrives at the scheduled time, we know that P must have halted. If no rocket arrives, we know that the operating system never had cause to launch it, and we conclude that P ran forever.

Hogarth refers to this hypercomputational system as an SAD_1 computer; it uses a standard Turing machine to run the underlying program P , but gains hypercomputational power from the geometrical properties of the spacetime in which that Turing machine finds itself. If we now adapt the construction to use a sequence of SAD_1 computers in an attempt to decide some question, the resulting (SAD_1 + singularity) system is called an SAD_2 machine, and so on. Finally, by dovetailing a sequence of machines, one from each level of the hierarchy, and sending the whole lot into an appropriate singularity, we obtain an AD machine. The SAD_n machines decide precisely those first order sentences which occupy the n th level of the Arithmetic Hierarchy, while the AD machine can decide the whole of arithmetic (Hogarth 2004).

The physicality of *Malament–Hogarth* spacetime is, however, debatable, since it clearly violates the Cosmic Censorship Hypothesis (Penrose 1999); in addition, there are various technical problems associated with the transmission of the successful-completion signal from computer to observer (Earman and Norton 1993, 1996). However, the related approach of Némethi and his colleagues (Némethi and Andr  ka 2006; Némethi and D  vid 2006) exploits instead the properties of slow-Kerr (i.e. massive slowly rotating uncharged) black holes, whence Cosmic Censorship is no longer an issue; they have, moreover, addressed the technical problems concerning signal transmission (Andr  ka et al. 2008) (see also the related paper, elsewhere this volume).

2.2.3 Quantum theoretical models

Quantum mechanics is, perhaps, mankind's most impressive scientific achievement to date; it enables us to predict various physical outcomes with remarkable accuracy

across a wide range of both everyday and exotic situations. In addition, as *It from Bit* demonstrates, there are clear parallels between quantum theory and information theory; since computation is largely seen as the study of information processing, it is not surprising that the field has proven fertile ground for researchers in both digital physics and hypercomputation theory.

One possible hypercomputational model in quantum theory is Kieu's adiabatic quantum algorithm for deciding Hilbert's Tenth problem, concerning the solution of Diophantine equations. Since this problem is known to be recursively undecidable (Matiyasevich 1993), Kieu's algorithm—essentially a method for searching infinite sets in finite time—must be hypercomputational. Although Kieu's claims are controversial and his algorithm has been disputed by various authors, he has sought to address these criticisms in a forthcoming paper (Kieu 2009). For the time being, therefore, the jury is out.

3 The standard path-integral formulation

As we explained in Sect. 2.2, we aim to reformulate the standard version of quantum theory from first principles in such a way that its computational aspects become essentially self-evident. We begin by recapitulating the (non-relativistic) path-integral formulation originally presented in Feynman (1948, Sects. 3–4; see also Feynman 1965). Given initial and final locations $q_I = (x_I, t_I)$ and $q_F = (x_F, t_F)$ (where $t_F > t_I$), the goal of the standard formulation is to determine the amplitude $\phi(q_F, q_I)$ that a particle P follows a trajectory $q_I \rightarrow q_F$ lying entirely within some prescribed non-empty open space-time region R . As Feynman shows, this amplitude can then be used to generate a Schrödinger wave-equation description of the system, whence this formulation is equivalent to other standard (non-relativistic) models of quantum theory. In Sect. 4, we will develop a generalised finitary formulation of the same amplitude, and show that it is equivalent to the standard path-integral formulation presented below.

For the sake of illustration, we shall assume that space is 1-dimensional, so that spatial locations can be specified by a single coordinate x —the extension to higher dimensions is straightforward. Furthermore, we shall assume in this paper that the region R is a simple rectangle of the form $R = X \times T$, where X and $T = (t_{\min}, t_{\max})$ are non-empty open intervals in \mathbb{R} ; this does not limit our results, because open rectangles form a base for the standard topology on \mathbb{R}^2 , and all of our formulae are derived via integration.³

Suppose, then, that a particle P is located initially at $q_I = (x_I, t_I)$, and subsequently at $q_F = (x_F, t_F)$, and that its trajectory from q_I to q_F is some continuous path lying entirely within the region $R = X \times T$. Choose some positive integer v , and split the duration $\delta t = t_F - t_I$ into $v + 1$ equal segments: for $n = 0, \dots, v + 1$, we define $t_n = t_I + n\delta t/(v + 1)$, so that $t_0 = t_I$ and $t_{v+1} = t_F$. We write x_0, \dots, x_{v+1} for the corresponding spatial locations, and define $q_n = (x_n, t_n)$. While each of the values x_n can vary from path to path, the values t_n are fixed. To distinguish this situation from the situation below (where t_n is allowed to vary), we shall typically write $q^\dagger = (x, t^\dagger)$ for those locations q_n whose associated t_n -value is fixed. We will also sometimes write $[q^\dagger]$ or $[q_1^\dagger, \dots, q_v^\dagger]$ for the

³ Integrating over a union of disjoint rectangles is the same as summing the component integrals: given any integrable function $f(x, t)$ defined on a disjoint union $R = \bigcup_x R_x$, we have $\int_R f = \sum_x \int_{R_x} f$.

arbitrary path $q_I = q_0^\dagger \rightarrow q_1^\dagger \rightarrow \cdots \rightarrow q_v^\dagger \rightarrow q_{v+1}^\dagger = q_F$. Apart from the fixed values $x_0 \equiv x_I$ and $x_{v+1} \equiv x_F$, each of the x_n is constrained only by the requirement that $x_n \in X$, whence the path $[q^\dagger]$ has v degrees of freedom.

In classical physics, the *action* associated with a path p is given by $S = \int_p L \, dt$, where the function $L = L(x(t), \dot{x}(t))$, the *Lagrangian*, is a function of position x and velocity \dot{x} , only. However, to form this integral we need to specify the motion of the particle in each subinterval $(t_n^\dagger, t_{n+1}^\dagger)$, so we assume that P follows some path $q_n^\dagger \rightarrow q_{n+1}^\dagger$ that is classically permissible. Each segment $q_n^\dagger \rightarrow q_{n+1}^\dagger$ of the path has associated classical action $S(q_{n+1}^\dagger, q_n^\dagger)$, and probability amplitude $\langle q_{n+1}^\dagger | q_n^\dagger \rangle$ defined for all q and (subsequent) q' by $\langle q' | q \rangle = \exp \{iS(q', q)/\hbar\}$. The action S is determined by the classical *Principle of Least Action*. This says that the classical path is one which minimises this action, so that $S(q', q) = \min \int_t^t L dt$. The total action associated with the path is $S[q_1^\dagger, \dots, q_v^\dagger] = \sum_n S(q_{n+1}^\dagger, q_n^\dagger)$ and the associated amplitude is the product $\langle q_F | q_v^\dagger \rangle \langle q_v^\dagger | q_{v-1}^\dagger \rangle \cdots \langle q_2^\dagger | q_1^\dagger \rangle \langle q_1^\dagger | q_I \rangle$. Summing over all such paths now yields the composite amplitude

$$\phi_v(q_F, q_I) = \frac{1}{A_v} \int \langle q_F | q_v^\dagger \rangle dx_v \langle q_v^\dagger | q_{v-1}^\dagger \rangle dx_{v-1} \cdots \langle q_2^\dagger | q_1^\dagger \rangle dx_1 \langle q_1^\dagger | q_I \rangle \quad (1)$$

where A_v is a normalisation factor. All that remains is to take the limit as $v \rightarrow \infty$, subject to the assumption that the resulting path $x = x(t)$ is continuous. This gives us the required amplitude $\phi(x_F, x_I)$ that the particle travels from q_I to q_F by a trajectory that lies entirely⁴ within R :

$$\phi(q_F, q_I) = \lim_{v \rightarrow \infty} \frac{1}{A_v} \int \langle q_F | q_v^\dagger \rangle dx_v \langle q_v^\dagger | q_{v-1}^\dagger \rangle dx_{v-1} \cdots \langle q_2^\dagger | q_1^\dagger \rangle dx_1 \langle q_1^\dagger | q_I \rangle.$$

4 A finitary formulation

In Sect. 3 we showed how the amplitude $\phi(q_F, q_I)$, that the particle P travels from q_I to q_F along some path lying entirely within the non-empty open spacetime region $R = X \times T$, is given by $\phi = \lim_{v \rightarrow \infty} \phi_v$. If we now write

$$\Delta_n = \phi_n - \phi_{n-1}, \quad (2)$$

it follows from the identity $\phi_v = (\phi_v - \phi_{v-1}) + \cdots + (\phi_1 - \phi_0) + \phi_0$ that

$$\lim_{v \rightarrow \infty} \phi_v = \lim_{v \rightarrow \infty} \left(\phi_0 + \sum_{n=1}^v \Delta_n \right) = \phi_0 + \sum_{n=1}^{\infty} \Delta_n.$$

This replacing of a limit with a sum is a key feature of our model, since it allows us to describe a system in terms of a set of mutually distinct finite sets of observations. We can think of this sum in terms of *correction factors*. For, suppose you were asked to estimate the amplitude $\phi(q_F, q_I)$ that some object or particle P will be observed at q_F , given that it had already been observed at q_I and was constrained to move within the region R . With no other information to hand, your best bet would be to assume that P follows some action-minimising classical path, and so the estimate you give is the associated amplitude $\langle q_F | q_I \rangle$.

⁴ Strictly, only the internal points of the trajectory are required to lie in R . Either (or both) of the endpoints q_I and q_F can lie outside R , provided they are on its boundary.

Some time later, you realise that one or more observations may have been made on the particle while it was moving from q_I to q_F , and that this would have perturbed the amplitude. To take account of these possibilities, you add a series of correction factors to your original estimate; first you add Δ_1 in case 1 observation had taken place, instead of the 0 observations you had originally assumed. Then you add Δ_2 in case there were actually 2 observations, and so on. Each Δ_n takes into account the extra information acquired by performing n observations instead of $n - 1$, and since the overall estimate needs to take all of the corrections into account, we have $\phi = \phi_0 + \sum \Delta_n$.

The simple truth, however, is that *continuous motion cannot be observed*, because making an observation takes time. The best we can ever do is to make a series of distinct measurements showing us where an object was at finitely many closely spaced instants t_1, t_2, \dots, t_v during the relocation from q_I to q_F . The classical spirit within us then tells us to extrapolate these discrete points into a continuous curve (namely, that path which ‘best’ joins the points). It is as if we draw the individual locations on celluloid, and then play a mental film projector to give ourselves the comfortable impression of continuous movement. But this mental film projector—represented in the standard formulation by the construction of $\lim \phi_v$ —is no part of physical observation; it represents instead an *assumption* about the way the world ‘ought to be’. All we can truthfully say is that the object was at such and such a location x_n when we observed it at time t_n , and was subsequently at location x_{n+1} at time t_{n+1} . Regardless of underlying reality (about which we can say virtually nothing), the *observed* universe is inherently discrete. We can ask ourselves how the motion appears if no observations are made; the composite answer, taking into account all potential observers, is given by some amplitude ψ_0 . If we ask how it appears if precisely v observations are made during the relocation from q_I to q_F , we get another amplitude ψ_v . Since these possibilities are all mutually exclusive, and account for every possible finitely observed relocation from q_I to q_F , the overall amplitude that the relocation happens is the sum of these amplitudes, namely some function $\psi = \sum \psi_v$.

Although they both involve infinite sums, these two descriptions are very different, because ψ_n tells us the amplitude for a path with a specific number of hops, while Δ_n describes what happens when we *change* the number of hops. Nonetheless, prompted by the formal structural similarity of the equations $\phi = \phi_0 + \sum_1^\infty \Delta_n$ and $\psi = \sum_0^\infty \psi_n$, we shall equate the two sets of terms, and attempt to find solutions. By requiring $\psi_0 = \phi_0$ and $\psi_n = \Delta_n$ for positive n , this will ensure that the description we generate—no matter how unnatural it might appear at first sight—satisfies $\phi = \psi$, whence it describes exactly the same version of physics as the standard formulation.

The surprising feature in what follows is that the description we generate is *not* unnatural. Quite the opposite. To see why, we need to remember that amplitudes are normally given in the form $\phi_n = \exp \{i(S_1 + \dots + S_n)/\hbar\}$. In very rough terms, we can think of the various S values as being essentially equal, so that $\phi_n \approx \exp \{inS/\hbar\}$. When we compute Δ_n , we are asking how ϕ_n changes when n changes; in other words, we can think of Δ_n in fairly loose terms as a measure of $d\phi_n/dn$. Again arguing loosely, we can calculate $d\phi_n/dn \approx iS\phi_n/\hbar$, and now it becomes clear why equating the two sets of terms works, for in essence, Δ_n is approximately proportional to ϕ_n . Since ψ_n is structurally similar to ϕ_n , in the sense that both measure the amplitude associated with a sequence of jumps, it is not surprising to find a similar relationship holding between Δ_n and ψ_n . Since the equations we form will eventually include integrals with normalisation factors, these factors will effectively absorb any remaining constants of proportionality.

4.1 Paths, actions and amplitudes

The standard formulation assumes that each trajectory $x(t)$ is a consistently future-pointing⁵ spacetime path; this is implicit in the continuity of the representation $x \equiv x(t)$, which assigns one location to each t in the interval $[t_I, t_F]$. Since our formulation rejects this assumption, we need to provide a different definition for *paths*.

We shall assume the abstract existence of a clock, represented by the integer variable τ , used to indicate the order in which observations occur. Each time the clock ticks, i.e. for each $\tau = 0, 1, 2, \dots$, the particle is observed to exist at some space-time location $q_\tau = (x_\tau, t_\tau)$. We call each transition $q_\tau \rightarrow q_{\tau+1}$ a *hop*. A finite sequence of consecutive hops $q_0 \rightarrow \dots \rightarrow q_{v+1}$ constitutes a *path*. As before, we take $q_0 = (x_I, t_I)$ and $q_{v+1} = (x_F, t_F)$, and consider the properties of an arbitrary path from q_I to q_F via v intermediate points, all of which are required to lie in the prescribed space-time region $R = X \times T$.

We again write $[q_1, \dots, q_v]$ for the path $q_I \rightarrow q_1 \rightarrow \dots \rightarrow q_v \rightarrow q_F$. However, whereas the intervals $t_{n+1} - t_n$ were formerly fixed to have identical duration $\frac{\delta t}{(v+1)}$, there is no constraint on the temporal separation $t_{\tau+1} - t_\tau$ in the finitary formulation; the path $q_0 \rightarrow \dots \rightarrow q_{v+1}$ therefore has $2v$ degrees of freedom, or *twice* the number in the standard formulation. Notice that we now write q_n rather than q_n^\dagger , to show that the value t_n is no longer fixed.

What is not clear at this stage is whether hops need necessarily always be future-pointing. The standard formulation forces this on us through its assumption that some continuous motion $t \mapsto x(t)$ is being observed, but this assumption is no longer relevant. We shall therefore describe two finitary formulations, one in which hops are unidirectional in time, and one in which space and time are treated symmetrically, in that hops can move both forwards and backwards in time as well as space. Both models are related to computation theory, but the second is by far the more interesting, both from a computational, and a physical, point of view. The mathematical distinction between the two models is minor. If time is unidirectional into the future, then $t_{\tau+1}$ must lie in the range $t_\tau < t_{\tau+1} \leq t_{\max}$. Otherwise, it can take any value in T .

In the standard formulation, any unobserved motion from one observation to the next is assumed to be classical, and its amplitude is determined by minimising the classical action S . Since we no longer assume that any such motion exists, we shall simply assume that each hop $q \rightarrow q'$ has a *hop amplitude*, denoted $\langle q'|q \rangle_h$, and that this amplitude (when it is non-zero) is associated with an abstract *hop action*, denoted $s_h(q', q)$, by the formula $\langle q'|q \rangle_h = e^{is_h(q', q)/\hbar}$. One of our tasks will be to identify the function s_h .

The amplitude associated with the path $[q_1, \dots, q_v]$ is defined, as usual, to be the product $\langle q_F|q_v \rangle_h \times \dots \times \langle q_1|q_I \rangle_h$. The amplitude computed by summing over all paths of this length will be denoted ψ_n , so that the overall *finitary amplitude* that the particle moves from q_I to q_F along a sequence of hops lying entirely within R is just $\psi(q_F, q_I) = \sum_{n=0}^{\infty} \psi_n$.

4.2 The finitary equations

Consider again the formulae giving the amplitude that a particle P follows a path from q_I to q_F that lies entirely within the region R , *subject to the assumption* that q_F occurs later than q_I —the standard formulation isn't defined when this isn't the case. We can write these in the form

⁵ As explained in his 1965 Nobel Prize address, Feynman 1965 subsequently described anti-particles as particles moving 'backwards in time'. In effect, our own approach adopts this temporal bi-directionality, and places it centre-stage.

$$\phi = \phi_0 + \sum_{n=1}^{\infty} \Delta_n \quad (3)$$

$$\psi = \psi_0 + \sum_{n=1}^{\infty} \psi_n \quad (4)$$

whence it is clear that one particular solution can be obtained by solving the infinite family of equations

$$\psi_0 = \phi_0 \quad (5)$$

$$\psi_n = \phi_n - \phi_{n-1} \quad (\text{i.e. } \psi_n = \Delta_n) \quad \text{for } n > 0 \quad (6)$$

to find the hop-action s_n . Since the terms ϕ_n and A_n are those of the standard formulation, we shall henceforth assume that S, ϕ_n, Δ_n and A_n are all *known functions* in what follows.

4.3 Solving the equations

As usual, we shall assume that q_F occurs later than q_I (so that $\phi_n = \phi_n(q_F, q_I)$ is defined for each n). We shall be careful to distinguish locations $q^\dagger = (x, t^\dagger)$ for which the time of observation is fixed in the standard formulation, from those of the form $q = (x, t)$ used in the finitary version, for which the value of t is variable. Note first that (1) can be rewritten to give us a recursive definition of ϕ_v , viz.

$$\begin{aligned} \phi_v(q_F, q_I) &= \frac{1}{A_v} \int \langle q_F | q_v^\dagger \rangle dx_v \langle q_v^\dagger | q_{v-1}^\dagger \rangle dx_{v-1} \cdots \langle q_2^\dagger | q_1^\dagger \rangle dx_1 \langle q_1^\dagger | q_I \rangle \\ &= \frac{A_{v-1}}{A_v} \int \langle q_F | q_v^\dagger \rangle dx_v \frac{1}{A_{v-1}} \int \langle q_v^\dagger | q_{v-1}^\dagger \rangle dx_{v-1} \cdots \langle q_2^\dagger | q_1^\dagger \rangle dx_1 \langle q_1^\dagger | q_I \rangle \\ &= \frac{A_{v-1}}{A_v} \int \langle q_F | q_v^\dagger \rangle \phi_{v-1}(q_v^\dagger, q_I) dx_v \end{aligned} \quad (7)$$

and an identical derivation gives ψ_v in the form

$$\psi_v(q_F, q_I) = \frac{B_{v-1}}{B_v} \int_X \int_{T'} \langle q_F | q_v \rangle_h \psi_{v-1}(q_v, q_I) dt_v dx_v \quad (8)$$

where the B_n are normalisation factors, and the integration range T' depends on whether we allow hops to jump backwards in time, or insist instead that they move only forwards (we consider the two cases separately, below).

Using (7) to substitute for ϕ_v in the definition (2) of Δ_n gives

$$\begin{aligned} \Delta_v(q_F, q_I) &= \phi_v(q_F, q_I) - \phi_{v-1}(q_F, q_I) \\ &= \left[\frac{A_{v-1}}{A_v} \int \langle q_F | q_v^\dagger \rangle \phi_{v-1}(q_v^\dagger, q_I) dx_v \right] - \phi_{v-1}(q_F, q_I). \end{aligned}$$

The case $v = 0$ is worth noting in detail. The amplitudes $\phi_0(q_F, q_I)$ and $\psi_0(q_F, q_I)$ describe the situation in which P moves from q_F to q_I without being observed. In the standard formulation, it is assumed in such circumstances that P follows some classical path for which the action S is minimal, while in the finitary formulation we assume that the particle *hops* directly from q_I to q_F . The amplitudes for these behaviours are $\langle q_F | q_I \rangle$ and $\langle q_F | q_I \rangle_h$, respectively. However, we need to remember that ϕ_0 and ψ_0 are defined in terms of their

contribution to the *overall* amplitudes ϕ and ψ ; it is important, therefore, to include the relevant normalisation factors. We therefore define, in accordance with (1), (3), (4) and (8),

$$\phi_0(q_F, q_I) = \frac{1}{A_0} \langle q_F | q_I \rangle \quad \text{and} \quad \psi_0(q_F, q_I) = \frac{1}{B_0} \langle q_F | q_I \rangle_h,$$

so that, whenever q_F occurs later than q_I ,

$$\langle q_F | q_I \rangle_h = \sigma \langle q_F | q_I \rangle \quad (9)$$

where

$$\sigma = B_0/A_0.$$

Taking principal logarithms on both sides of (9) now gives

$$s_h(q_F, q_I) = S(q_F, q_I) - i\hbar \log \sigma$$

and if we assume that s_h should be real-valued (the classical action S is always real-valued), then $\log \sigma$ must be a real multiple of i , say $\sigma = e^{i\rho}$ where $\rho \in \mathbb{R}$, whence $|\sigma|^2 = 1$. Consequently, $|\langle q_F | q_I \rangle_h|^2 = |\langle q_F | q_I \rangle|^2$, and the two formulations assign the same standard and finitary probabilities to the relocation $q_I \rightarrow q_F$, whenever this is unobserved and future-directed. Moreover, since

$$s_h(q_F, q_I) = S(q_F, q_I) + \rho\hbar$$

we see that our earlier intuition is essentially confirmed: the hop-action s_h (the best estimate of the path-amplitude, given that no observations will be made) is just the classical action S , though possibly re-scaled by the addition of a constant action of size $\rho\hbar$ (which we can think of as a kind of ‘zero-point’ action). For the purposes of this paper, the values of ρ and $\sigma = e^{i\rho}$ are essentially arbitrary; we shall leave ρ (and hence σ) an undetermined parameter of the model, in terms of which

$$B_0 = \sigma A_0 \quad (10)$$

and

$$s_h(q_F, q_I) = S(q_F, q_I) + \rho\hbar \quad \text{if } q_F \text{ occurs after } q_I. \quad (11)$$

The physical significance of ρ is discussed briefly in Sect. 4.5, in relation to *null-hops*.

4.4 The unidirectional model

If we wish to allow only future-pointing hops—we shall call this the *unidirectional* model—there is little left to do. We know from (5) and (6) that each function ψ_n is defined in terms of the known functions ϕ_0 and Δ_n . It only remains to identify the hop amplitude s_h and the normalisation factors B_n . As explained above, our solutions will be given in terms of the undetermined phase parameter σ .

Since the side-condition on (11) is satisfied, the hop amplitude is given in terms of the classical action by the formula $\langle q' | q \rangle_h = \sigma \langle q' | q \rangle = \sigma \exp\{iS(q', q)/\hbar\}$, whenever q' follows q .

To find the normalisation factors, we note first that (10) gives us the value $B_0 = \sigma A_0$ directly. Next, when $v > 0$, we observe that, since t_v must come after t_{v-1} , the range T' in (8) is the interval (t_{v-1}, t_F) . Consequently,

$$\begin{aligned}
 \psi_v(q_F, q_I) &= \frac{B_{v-1}}{B_v} \int_X \int_{t_{v-1}}^{t_F} \langle q_F | q_v \rangle_h \psi_{v-1}(q_v, q_I) dt_v dx_v \\
 &= \frac{\sigma B_{v-1}}{B_v} \int_X \int_{t_{v-1}}^{t_F} \langle q_F | q_v \rangle \psi_{v-1}(q_v, q_I) dt_v dx_v
 \end{aligned} \tag{12}$$

When $v = 1$, (4) can be rewritten

$$\begin{aligned}
 \psi_1(q_F, q_I) &= \frac{\sigma B_0}{B_1} \int_X \int_{t_I}^{t_F} \langle q_F | q_1 \rangle \psi_0(q_1, q_I) dt_1 dx_1 \\
 &= \frac{\sigma B_0}{B_1} \int_X \int_{t_I}^{t_F} \langle q_F | q_1 \rangle \frac{1}{B_0} \langle q_1 | q_I \rangle_h dt_1 dx_1 \\
 &= \frac{\sigma^2}{B_1} \int_X \int_{t_I}^{t_F} \langle q_F | q_1 \rangle \langle q_1 | q_I \rangle dt_1 dx_1
 \end{aligned}$$

and, since $\psi_1 = \Delta_1$, this gives us

$$B_1 = \left(\frac{\int_X \int_{t_I}^{t_F} \langle q_F | q_1 \rangle \langle q_1 | q_I \rangle dt_1 dx_1}{\Delta_1(q_F, q_I)} \right) \sigma^2.$$

Finally, for $v > 1$, (12) becomes

$$\begin{aligned}
 \Delta_v(q_F, q_I) &= \psi_v(q_F, q_I) \\
 &= \frac{\sigma B_{v-1}}{B_v} \int_X \int_{t_{v-1}}^{t_F} \langle q_F | q_v \rangle \psi_{v-1}(q_v, q_I) dt_v dx_v \\
 &= \frac{\sigma B_{v-1}}{B_v} \int_X \int_{t_{v-1}}^{t_F} \langle q_F | q_v \rangle \Delta_{v-1}(q_v, q_I) dt_v dx_v
 \end{aligned}$$

and hence B_v can be defined recursively, as

$$B_v = \frac{\sigma B_{v-1}}{\Delta_v(q_F, q_I)} \int_X \int_{t_{v-1}}^{t_F} \langle q_F | q_v \rangle \Delta_{v-1}(q_v, q_I) dt_v dx_v$$

4.5 The bidirectional model

Far more interesting is the case where hops are allowed to jump backwards as well as forwards in time. It is important to note that the derivation of B_v given above for the unidirectional model no longer works, because it relies on using (9) to replace $\langle q_F | q_v \rangle_h$ with $\sigma \langle q_F | q_v \rangle$, and on (6) to replace $\psi_{n+1}(q_v, q_I)$ with $\Delta_{n+1}(q_v, q_I)$. But our use of (9) assumes that q_F occurs after q_v , and that of (6) that q_v comes after q_I , and neither assumption is generally valid in the bidirectional model. Consequently, before we can

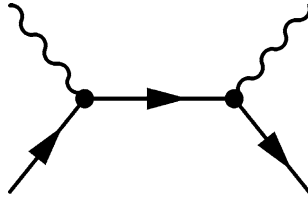


Fig. 1 Anti-matter can be thought of as matter moving backwards in time. A particle arrives at bottom left, and the corresponding antiparticle (shown as usual with the *arrow* reversed) at bottom right; they annihilate to produce two gamma rays, emitted top left and top right. Time advances up the page

make progress, we need to decide how $\langle q'|q \rangle_h$ should be defined when the hop $q \rightarrow q'$ moves *backwards* in time.

To address this problem, we recall the standard interpretation of *anti-matter* as ‘matter moving backwards in time’. For example, the Feynman diagram in Fig. 1 shows how the annihilation of e.g. an electron and a positron (its antiparticle) to form two photons can be interpreted instead as showing an electron that moves forward in time, interacts with the photons, and then returns into the past.

Accordingly, whenever we are presented with a backwards hop by the particle P , we re-interpret it as a *forwards* hop by the appropriate anti-particle, \bar{P} . Writing \bar{S} for the classical action associated with the antiparticle \bar{P} , we therefore define

$$s_h(q_F, q_I) = \begin{cases} \rho\hbar + S(q_F, q_I) & \text{if } q_I \text{ is earlier than } q_F, \text{ and} \\ \rho\hbar + \bar{S}(q_I, q_F) & \text{if } q_I \text{ is later than } q_F. \end{cases} \quad (13)$$

It is tempting to assume that \bar{S} is just the negative of S , but this need not be the case. For example, since photons are their own anti-particles, they would require $\bar{S} = S$. Or consider an electron moving in both an electric and a gravitational field. If we replaced it with a positron, the electric forces would reverse, but the gravitational forces would remain unchanged, and the overall change in action would reflect both effects.

4.5.1 Spatial hops—the physical meaning of σ

What about purely spatial hops that move the particle P sideways in space, without changing its temporal coordinate? There are two cases to consider. If $q_F = q_I$, the particle has not actually moved, and the classical solution $S(q, q) = 0$ holds valid. Consequently, we can simply extend our existing solution by defining $s_h(q, q) = \rho\hbar$, or $\langle q|q \rangle_h = \sigma$. This, then, explains the physical significance of σ —it is the amplitude associated with the *null hop*, i.e. that hop which leaves the particle in its original location from one observation to the next.

If q_F and q_I differ in their x (but not their t) values, we shall simply take $\langle q_F|q_I \rangle_h = 0$; i.e. we ban all such hops (this definition is, of course, purely arbitrary, and other definitions may be more appropriate in regard to other investigations⁶; but for our current purposes the specific choice of purely spatial hop action makes little difference, because the paths in question contribute nothing to the integrals we shall be constructing). This doesn’t mean,

⁶ For example, suppose we know (from wave-equation methods, say) that P has amplitude $\eta(x)$ to be at location $x^\dagger = (x, t^\dagger)$, for each $x \in X$. A more intuitive solution might then be to take $\langle x^\dagger|y^\dagger \rangle_h = \eta(x^\dagger)/\eta(y^\dagger)$. This gives $\langle x^\dagger|x^\dagger \rangle_h = 1$ in agreement with the ‘classical amplitude’, but also provides information about the relative amplitudes of all other spatial locations at time t^\dagger .

of course, that a path cannot be found from q_I to a simultaneous location q_F —it can, via any past or future location—but that more than one hop is required to complete the journey. Indeed, the possibility of purely spatial relocations is highly significant, since one could interpret them as explaining quantum uncertainty: one cannot say definitely where a particle is at any given time t , precisely *because* it is able to relocate from one location to another, with no overall change in t .

4.5.2 Solving the equations

As before, we know from (5) and (6) that each function ψ_n is defined in terms of the known functions ϕ_0 and Δ_n , and it remains to identify the hop amplitude s_h and the normalisation factors B_n . Once again, our solutions will be given in terms of the undetermined phase parameter σ . As always, we assume that $t_I < t_F$, although we allow individual hops to move backwards through time.

To define the hop amplitude, we appeal to (13), and the relationship $\langle q'|q \rangle_h = e^{is_h(q',q)/\hbar}$. Taken together with our discussion of spatial hops, these allow us to define s_h fully:

$$\langle q_F|q_I \rangle = \begin{cases} \sigma \overline{\langle q_I|q_F \rangle} & \text{if } q_F \text{ is earlier than } q_I, \\ \sigma \langle q_F|q_I \rangle & \text{if } q_F \text{ is later than } q_I \\ \sigma & \text{if } q_F = q_I, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

where $\overline{\langle q_I|q_F \rangle} = \exp\{i\bar{S}(q_I, q_F)/\hbar\}$ is the ‘classical amplitude’ associated with the antiparticle. This idea extends throughout the functions defined in this section; for example, when q' is earlier than q , we write $\overline{\psi}_n(q, q')$ for the amplitude that the antiparticle follows some path $q' \rightarrow q$ lying entirely within R . We will see below that the amplitude functions $\psi_n(q', q)$ and $\overline{\psi}_n(q', q)$ are, as one would expect, related to one another in a mutually recursive way.

Now we consider the normalisation constants B_n . We already know that $B_0 = \sigma A_0$, so we consider the case when $n > 0$. Because hops are allowed to move in both directions through time, the integration range T' in (8) is the whole of T . Consequently, (8) becomes

$$\psi_v(q_F, q_I) = \frac{B_{v-1}}{B_v} \int_X \int_T \langle q_F|q_v \rangle_h \psi_{v-1}(q_v, q_I) dt_v dx_v.$$

The integral over T splits into three parts, depending on the value of t_v relative to t_I and t_F . We have

$$\begin{aligned} \psi_v(q_F, q_I) &= \frac{B_{v-1}}{B_v} \int_X \int_T \langle q_F|q_v \rangle_h \psi_{v-1}(q_v, q_I) dt_v dx_v \\ &= \frac{B_{v-1}}{B_v} \int_X [I_L(x_v) + I_M(x_v) + I_R(x_v)] dx_v \end{aligned} \quad (14)$$

where $I_L(x_v)$ is the integral over $[t_{\min}, t_I]$, $I_M(x_v)$ over $[t_I, t_F]$ and $I_R(x_v)$ over $[t_F, t_{\max}]$.

When $v = 1$, (14) becomes

$$\psi_1(q_F, q_I) = \frac{B_0}{B_1} \int_X [I_L(x_1) + I_M(x_1) + I_R(x_1)] dx_1$$

and the integrals I_L , I_M and I_R are defined by

$$\begin{aligned}
I_L(x_1) &= \sigma \int_{t_{\min}}^{t_I} \langle q_F | q_1 \rangle_h \psi_0(q_1, q_I) \, dt_1 = \frac{\sigma^2}{B_0} \int_{t_{\min}}^{t_I} \langle q_F | q_1 \rangle \overline{\langle q_I | q_1 \rangle} \, dt_1 \\
I_M(x_1) &= \sigma \int_{t_I}^{t_F} \langle q_F | q_1 \rangle_h \psi_0(q_1, q_I) \, dt_1 = \frac{\sigma^2}{B_0} \int_{t_I}^{t_F} \langle q_F | q_1 \rangle \langle q_1 | q_I \rangle \, dt_1 \\
I_R(x_1) &= \sigma \int_{t_F}^{t_{\max}} \langle q_F | q_1 \rangle_h \psi_0(q_1, q_I) \, dt_1 = \frac{\sigma^2}{B_0} \int_{t_F}^{t_{\max}} \overline{\langle q_1 | q_F \rangle} \langle q_1 | q_I \rangle \, dt_1.
\end{aligned}$$

Thus $I_L(x_1) + I_M(x_1) + I_R(x_1) =$

$$\frac{\sigma^2}{B_0} \left[\int_{t_{\min}}^{t_I} \langle q_F | q_1 \rangle \overline{\langle q_I | q_1 \rangle} + \int_{t_I}^{t_F} \langle q_F | q_1 \rangle \langle q_1 | q_I \rangle + \int_{t_F}^{t_{\max}} \overline{\langle q_1 | q_F \rangle} \langle q_1 | q_I \rangle \right]$$

and $\psi_1(q_F, q_I)$ equals

$$\frac{\sigma^2}{B_1} \left[\int_{t_{\min}}^{t_I} \langle q_F | q_1 \rangle \overline{\langle q_I | q_1 \rangle} + \int_{t_I}^{t_F} \langle q_F | q_1 \rangle \langle q_1 | q_I \rangle + \int_{t_F}^{t_{\max}} \overline{\langle q_1 | q_F \rangle} \langle q_1 | q_I \rangle \right].$$

On the other hand, (2) tells us that $\psi_1 = \Delta_1$, and so B_1 equals

$$\frac{\sigma^2}{\Delta_1(q_F, q_I)} \times \left[\int_{t_{\min}}^{t_I} \langle q_F | q_1 \rangle \overline{\langle q_I | q_1 \rangle} + \int_{t_I}^{t_F} \langle q_F | q_1 \rangle \langle q_1 | q_I \rangle + \int_{t_F}^{t_{\max}} \overline{\langle q_1 | q_F \rangle} \langle q_1 | q_I \rangle \right].$$

Finally, when $v > 1$, the integrals I_L , I_M and I_R are given by

- $I_L(x_v) = \sigma \int_{t_{\min}}^{t_I} \langle q_F | q_v \rangle \overline{\Delta_{v-1}(q_I, q_v)} \, dt_v;$
- $I_M(x_v) = \sigma \int_{t_I}^{t_F} \langle q_F | q_v \rangle \Delta_{v-1}(q_v, q_I) \, dt_v;$
- $I_R(x_v) = \sigma \int_{t_F}^{t_{\max}} \overline{\langle q_v | q_F \rangle} \Delta_{v-1}(q_v, q_I) \, dt_v,$

and (14) gives us B_v recursively,

$$\begin{aligned}
B_v &= \frac{\sigma B_{v-1}}{\Delta_v(q_F, q_I)} \int_X \left\{ \int_{t_{\min}}^{t_I} \langle q_F | q_v \rangle \overline{\Delta_{v-1}(q_I, q_v)} \, dt_v \right. \\
&\quad + \int_{t_I}^{t_F} \langle q_F | q_v \rangle \Delta_{v-1}(q_v, q_I) \, dt_v \\
&\quad \left. + \int_{t_F}^{t_{\max}} \overline{\langle q_v | q_F \rangle} \Delta_{v-1}(q_v, q_I) \, dt_v \right\}.
\end{aligned}$$

5 Computational interpretation of the model

To illustrate the full computational significance of our reformulation (especially the bidirectional version), we first need to digress slightly, and explain Eilenberg's (1974)

X-machine model of computation. This is an extremely powerful computational model, which easily captures (and extends) the power of the Turing machine. We will then show that a particle's trajectory can be regarded as an *X-machine* drawn in spacetime, and that (a minor variant of) this machine computes its own amplitude (as a trajectory).

5.1 *X-machines*

An *X-machine* $M = F^\Lambda$ (where X is a data type) is a finite state machine F over some alphabet A , together with a *labelling* function $\Lambda: a \mapsto a^\Lambda: A \rightarrow R(X)$, where $R(X)$ is the ring of relations of type $X \leftrightarrow X$.

Each word $w = a_1, \dots, a_n$ in the language $|F|$ recognised by the machine F can be transformed by Λ into a relation w^Λ on X , using the scheme

$$w^\Lambda = a_1^\Lambda \circ \dots \circ a_n^\Lambda$$

and taking the union of these relations gives the relation $|F^\Lambda|$ computed by the machine,

$$|F^\Lambda| = \bigcup \{w^\Lambda \mid w \in |F|\}.$$

If we want to model a relation of type $Y \leftrightarrow Z$, for data types $Y \neq Z$, we equip the machine with encoding and decoding relations, $E: Y \rightarrow X$ and $D: X \rightarrow Z$. Then the behaviour computed by the extended machine is the relation $E \circ |F^\Lambda| \circ D$.

Although the language $|F|$ is necessarily regular, the computational power of the *X-machine* model is unlimited. For, given any set-theoretic relation $\zeta: Y \rightarrow Z$, we can compute it using the trivial (2-state, 1-transition)-machine with $X = Y \times Z$, by picking any $z^\dagger \in Z$, and using the encoder $y^E = (y, z^\dagger)$, the decoder $(y, z)^D = z$, and labelling $a^\Lambda = \tilde{\zeta}$, where $(y, z^\dagger)^\zeta = (y, \zeta(y))$. For now, given any $y \in Y$, we have $|F^\Lambda| = \bigcup \{a^\Lambda\} = \tilde{\zeta}$, and

$$y^{(E \circ |F^\Lambda| \circ D)} = y^{(E \circ \tilde{\zeta} \circ D)} = (y, z^\dagger)^{(\tilde{\zeta} \circ D)} = \bigcup (y, \zeta(y))^D = \zeta(y).$$

5.2 Computation by admissible machines

In our case, all of the path relations we consider will be constant multipliers of the form $k_c: z \mapsto zc$, where $c, z \in \mathbb{C}$. The resulting machine behaviour will therefore be a set of such multipliers, and we can meaningfully form their sum (which is again a multiplier). For reasons that will shortly become clear, however, we will restrict attention to those paths which visit each state of the machine at least once. We therefore define the *additive behaviour* of such a machine $M = F^\Lambda$ to be the function $|M|^+$ on \mathbb{C} given by

$$|M|^+(z) = \sum \{w^\Lambda(z) \mid w \in |F|, w \text{ visits each state of } F \text{ at least once}\}.$$

If M is a machine of this form, we will declare the behaviour of M to be the function $|M|^+$, and speak of M as an *additive X-machine*. Any finitary path $[q] = q_I \rightarrow q_1 \rightarrow \dots \rightarrow q_v \rightarrow q_F$ generates an additive *X-machine* M_q with state set $\{q_I, q_1, \dots, q_v, q_F\}$, alphabet $A = \{h_0, \dots, h_v\}$, and transitions $\{q_n \xrightarrow{h_n} q_{n+1} \mid n = 0, \dots, v\}$. Each transition in the machine is a hop along the path, and is naturally associated with the function $h_n^\Lambda = \lambda z \cdot (z \cdot \langle q_{n+1} | q_n \rangle_h) : \mathbb{C} \rightarrow \mathbb{C}$ that multiplies any input amplitude z by the hop amplitude $\langle q_{n+1} | q_n \rangle_h$. If M_q is an additive *X-machine* generated by some path $[q]$ with initial state q_I , final state q_F , and intermediate states in R , we shall say that M is *admissible*, and

that $[q]$ generates M . We claim that each path computes its own amplitude, when considered as the machine it generates.

5.2.1 Computation by the unidirectional model

For unidirectional machines, each hop h_n involves a jump forward in time, so the states $\{q_n\}$ must all be distinct, and the path $[q]$ forms a future-pointing chain through spacetime. Consequently, the machine M_q recognises precisely one string, and the additive and standard behaviours of the X -machine are identical. The function computed by this path maps each $z \in \mathbb{C}$ to

$$z^{[(h_0^\Lambda) \circ \dots \circ (h_n^\Lambda)]} = z \times \langle q_{n+1} | q_n \rangle_h \langle q_n | q_{n-1} \rangle_h \cdots \langle q_1 | q_0 \rangle_h = z \times \psi[q]. \quad (15)$$

As claimed, therefore, each (unidirectional) trajectory directly computes its own contribution to the amplitude of any path containing it.

5.2.2 Computation by the bidirectional model

Equation (15) holds also for unidirectional paths in bidirectional machines, but the general physical interpretation is more complicated, because of the possibility of loops. Essentially, we need to distinguish carefully between two related questions, viz.

- what is the amplitude that the path $[q]$ is traversed?
- what is the amplitude that the path $[q]$ is *observed* to have been traversed?

To see why, let us suppose that the path $[q]$ contains only one loop, and that m is minimal such that $q_{m+1} = q_{n+1}$ for some n satisfying $m < n$; write the associated sequence of hops as a concatenation of three segments, viz. $h_0 \dots h_n = u.v.w$, where $u = h_0 \dots h_m$, $v = h_{m+1} \dots h_n$ and $w = h_{n+1} \dots h_v$. Since v represents a spacetime loop from q_{m+1} back to $q_{n+1} = q_{m+1}$, there is no observable difference between any of the paths $u.v^j.w$, for $j \geq 1$. Consequently, while the amplitude for the path $[q]$ is just $\psi[q]$, the amplitude that this path is *observed* is instead the amplitude $\psi^*[q] = \sum_{j=1}^{\infty} \psi[u] \times (\psi[v])^j \times \psi[w]$.

More generally, given the machine F generated by any bidirectional trajectory $[q]$, and any two strings α, β which are recognised by F , and which visit each state at least once, there will be no observable difference between α and β . Consequently, if we define

$$F^+ = \{w^\Lambda | w \in |F|, w \text{ visits each state at least once}\}$$

then the amplitude ψ^+ that $[q]$ is *observed* to have been the path traversed will satisfy, for $z \in \mathbb{C}$,

$$z.\psi^+ = \sum \{w^\Lambda(z) | w \in F^+\} = |F^\Lambda|^+(z)$$

and once again, if we think of $[q]$ as an additive X -machine, it computes its own contribution to the amplitude of any path containing it.

6 Concluding arguments

Recall that an additive X -machine M is *admissible* provided there is some finitary bidirectional path $[q]$ that generates it. Say that two paths $[q]_1$ and $[q]_2$ are *equivalent*, provided

they generate precisely the same admissible machine M . Clearly, this *is* an equivalence relation, and given any path $[q]$, there will some equivalence class \tilde{q} containing it. Moreover, the amplitude $|M|^+$ is given by summing the amplitudes of the various paths in \tilde{q} . Consequently, summing over all paths is the same as summing over all admissible machines, so that (regarding $\psi(q_F, q_I)$ as a multiplier),

$$\psi(q_F, q_I) = \sum \{|M|^+ |M \text{ is admissible}\},$$

and $\psi(q_F, q_I)$ can be regarded as integrating all of the admissible machine amplitudes. In the bidirectional formulation, then, the nature of motion in quantum theory reveals itself to be inherently computational. It is not that trajectories can be computed; rather, they *are* computations. As a particle hops through spacetime, it simultaneously *constructs* and *executes* a computational state machine, and the amplitude computed by this machine is precisely the amplitude of the trajectory that constructed it.

In Sect. 2.1, we noted how digital physics assumes the existence of a computation that computes each universe's history, which suggests that the 'computer' which executes the computation is somehow external to the universes being constructed. In contrast, the bidirectional model is telling us that each universe is a *process*, in which each trajectory is a sub-process which computes its own amplitude. Moreover, all of these sub-processes interact with one another non-locally, because hop amplitudes are based on the classical action, and this in turn depends on the ever-changing spacetime distribution of the other particles. In other words, as we have argued elsewhere, quantum theory is best thought of, not in terms of computation, but in terms of *interactive formal processes* (Stannett 2007).

Clearly, this idea has echoes of *It from Bit*, and indeed the bidirectional model helps explain Wheeler's delayed-choice experiment. The apparent paradox relies on two assumptions concerning the experimental set-up. First, the photon must pass through the barrier in order to be observed on the other side; and second, we can reliably identify a time by which the photon has travelled beyond the barrier (we need to make our delayed choice after this time). Both of our reformulations refute the first assumption (the discontinuous nature of hop-based motion means that the Intermediate Value Theorem cannot be invoked to prove that the trajectory necessarily passes through the barrier), while the bidirectional model also refutes the second assumption, since there is no reliable sense in which the decision can be said to have been made 'after' the trajectory intersects the barrier. Thus the delayed-choice experiment contains no paradox, and there is nothing to explain.

We should also be clear as to what our reformulation does *not* say. Throughout this discussion we have focussed on the computational nature of trajectories, but it should be stressed that there is an important distinction to be drawn between what a process *does*, and how that process is *structured*. This is the same distinction as that highlighted in Sect. 2.1 between Schmidhuber's and Tegmark's versions of the computational universe hypothesis: whereas Schmidhuber considers process evolutions to be computable, Tegmark requires instead that their descriptions be computable. In our case, while we know that each trajectory computes its amplitude, we cannot say that the amplitude itself is necessarily 'computable' in the Turing sense, because we cannot as yet identify the extent to which the two forms of computation are related. As a *process*, each trajectory is computational, but the *values* it manipulates need not be.

7 Open questions

- (a) Clearly, we need to determine the relationship between trajectory computations and Turing computations. There must certainly be some such relationship, because the admissible *X*-machine model underpinning trajectory computation is closely related to the Finite State Machine, which in turn underpins the basic structure of the Turing machine. Are values (like the processes that generate them) constrained to be computable in any standard sense?
- (b) Although we have exchanged continuous motion for motion based on discrete hops, we have not as yet done away with continuous spaces in their entirety, because many of the expressions given in this paper make use of integration. As we argued above, continuity is not directly observable, so we would prefer a purely discrete model. We should therefore investigate the extent to which the formulation presented here can be re-expressed in purely formal terms, for example using the π -calculus (a standard theoretical vehicle for modelling mobile distributed process-based systems) (Milner 1999; Sangiorgi and Walker 2001). More straightforwardly, can we adapt the models presented here—for example, by replacing integrals with sums—to generate a truly *discrete* models of physics?
- (c) Suppose we impose the condition that whenever a particle hops inside some arbitrary region (which we can think of as the interior of an event horizon), it cannot hop back out again. This will have a global influence upon trajectory amplitudes in the bidirectional model, because every journey would otherwise have had the option to include hops that pass through the excluded region. In particular, the observed positions of geodesics (assuming these can be modelled in terms of finite trajectories?) can be expected to change position, whence the presence of the excluded region will generate a perceived ‘warping’ of spacetime geometry. Does this warping agree with the warping predicted by, e.g. general relativity? Can the bidirectional model be extended to give a model of quantum gravity?
- (d) Feynman’s original path-integral methods appear to make various assumptions which we have rejected, including such mainstays of real-world observation as the *arrow of time* and the *continuity of motion*. The status of these assumptions in Feynman’s formulation needs, therefore, to be considered in more depth than has been possible here. It may be that they are spurious elements of his construction which play no actual rôle, and which are therefore logically independent of his formulae. But if they do indeed play a relevant part in his formulation, they must necessarily become *provable theorems* within both the unidirectional and bidirectional models presented here, because our models agree with Feynman’s *by construction*. That is, any property that is (a) expressible in terms of ‘what is seen by observers’, and (b) ‘built-into’ Feynman’s assumptions, must necessarily reappear from our own equations, since these give identical results when used to calculate amplitudes.

Acknowledgements This research was supported in part by the EPSRC HyperNet project (Hypercomputation Research Network, grant number EP/E064183/1).

References

- Andréka H, Németi I, Németi P (2008) General relativistic hypercomputing and foundation of mathematics. In: Calude CS, Costa JF (eds) *Physics and computation (renaissance) international workshop*, Vienna,

- Austria, August 25–28, 2008: Pre-proceedings, vol 327 of CDMTCS Research Report Series, pp 210–229, University of Auckland, New Zealand, July 2008. Centre for Discrete Mathematics and Theoretical Computer Science
- Beggs E, Costa JF, Loff B, Tucker JV (2008) Computational complexity with experiments as oracles. *Proc R Soc A* 464:2777–2801
- Bekenstein JD (1972) Black holes and the second law. *Lettere al Nuovo Cimento* 4:737
- Bekenstein JD (1973) Black holes and entropy. *Phys Rev D* 7:2333
- Copeland BJ (2002) The Church-Turing Thesis. In: Zalta EN (ed) *The Stanford encyclopedia of philosophy*. Stanford University, Stanford, CA, Fall 2002. Online publication, <http://plato.stanford.edu/archives/fall2002/entries/church-turing/>. Accessed 4 July 2008
- Deutsch D (1985) Quantum theory, the Church-Turing principle and the universal quantum computer. *Proc R Soc Ser A* 400:97–117
- Earman J, Norton J (1993) Forever is a day: supertasks in Pitowsky and Malament–Hogarth spacetimes. *Philos Sci* 5:22–42
- Earman J, Norton J (1996) Infinite pains: the trouble with supertasks. In: Morton A, Stich S (eds) *Benacerraf and his critics*. Blackwell, Cambridge, pp 231–261
- Eilenberg S (1974) *Automata, languages, and machines*, vol A. Academic Press, London
- Etesi G, Némethi I (2002) Non-Turing computability via Malament–Hogarth space-times. *Int J Theor Phys* 41(2):341–370
- Feynman RP (1948) Space–time approach to non-relativistic quantum mechanics. *Rev Mod Phys* 20(2):367–387
- Feynman RP (1965) The development of the space-time view of quantum electrodynamics. In: *Les Prix Nobel 1965*, Stockholm. Imprimerie Royale P.A. Norstedt & Soner, pp 172–191
- Gandy R (1980) Church’s thesis and principles for mechanisms. In: Barwise J, Keisler HJ, Kunen K (eds) *The Kleene symposium*. North-Holland, Amsterdam, pp 123–148
- Goldreich O (2008) *Computational complexity: a conceptual perspective*. Cambridge University Press, Cambridge
- Hawking SW (1974) Black hole explosions? *Nature* 248:30–31
- Hogarth M (1992) Does general relativity allow an observer to view an eternity in a finite time? *Found Phys Lett* 5:173–181
- Hogarth M (2004) Deciding arithmetic using *SAD* computers. *Br J Philos Sci* 55(4):681–691
- Horgan J (1991) Profile: physicist John A. Wheeler, Questioning the “It from Bit”. *Scientific American*, June 1991, pp 36, 37
- Kieu T (2009) Hypercomputability of quantum adiabatic processes. *Int J Unconv Comput (Special Issue on Hypercomputation)* 4 (In press)
- Kleene SC (1952) *Introduction to metamathematics*. North Holland, Amsterdam
- Lloyd S (2006) A theory of quantum gravity based on quantum computation. Preprint, arXiv:quant-ph/0501135. Accessed 4 July 2008
- Matiyasevich Y (1993) *Hilbert’s tenth problem*. MIT Press, Cambridge, MA
- Milner R (1999) *Communicating and mobile systems: the π -calculus*. Cambridge University Press, Cambridge
- Némethi I, Andr  ka H (2006) Can general relativistic computers break the Turing barrier? In: Beckmann A, Berger U, L  we B, Tucker JV (eds) *Logical approaches to computational barriers (Proc Conf CiE 2006, Swansea, UK, July 2006)*, vol 3988 of *Lecture notes in computer science*, Springer-Verlag, Berlin, pp 398–412
- N  methi I, D  vid Gy (2006) Relativistic computers and the Turing barrier. *J Appl Math Comput* 178:118–142
- Penrose R (1990) *The Emperor’s new mind*. Oxford University Press, Oxford
- Penrose R (1999) The question of cosmic censorship. In: Wald RM (ed) *Black holes and relativistic stars*. University of Chicago Press, Chicago, IL. Reprinted in *J Astrophys Astron* 20:233–248 (1998)
- Sangiorgi D, Walker D (2001) *The π calculus: a theory of mobile processes*. Cambridge University Press, Cambridge
- Schmidhuber J (1997) A computer scientist’s view of life, the universe, and everything. In: Freksa C (ed) *Foundations of computer science: potential–Theory–Cognition*, vol 1337 of *Lecture notes in computer science*, Springer, Berlin, pp 201–208
- Schmidhuber J (2000) Algorithmic theories of everything. Technical Report IDSIA–20–00, IDSIA, Galleria 2, 6928 Manno (Lugano), Switzerland, December 2000. Available Online, arXiv:quant-ph/0011122. Accessed 4 July 2008
- Shor P (1997) Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. In: Goldwasser S (ed) *Proceedings of 35th annual symposium on foundations of computer science*. IEEE Computer Society Press, Los Alamitos, CA, p 124, 1994. Revised version (1995),

- available Online, arXiv:quant-ph/9508027v2. Accessed 4 July 2008. Subsequently published as SIAM J Sci Stat Comput 26:1484
- Stannett M (2003) Computation and hypercomputation. *Minds Mach* 13(1):115–153
- Stannett M (2006) The case for hypercomputation. *Appl Math Comput* 178:8–24
- Stannett M (2007) Physical hypercomputation. In: Adamatzky A, Bull L, De Lacy Costello B, Stepney S, Teuscher C (eds) *Unconventional computing*. Luniver Press, Beckington, UK, pp 265–288
- Susskind L, Lindesay J (2005) An introduction to black holes, information and the string theory revolution—the holographic universe. World Scientific, Singapore
- Tegmark M (2008) The mathematical universe. *Found Phys* 38(2):101–150. Preprint available Online, arXiv:0704.0646 [gr-qc]. Accessed 4 July 2008
- Wheeler J (1990) Recent thinking about the nature of the physical world. It from Bit. In: Mendell RB, Mincer AI (eds) *Frontiers in cosmic physics; symposium in memory of Serge Alexander Korff*, held on September 13, 1990, in New York, NY, vol 655 of *Annals of the New York Academy of Sciences*. Academy of Sciences, New York, NY, p 349
- Wolfram S (2002) *A new kind of science*. Wolfram Media, Champaign, IL. ISBN 1-57955-008-8
- Xia Z (1992) The existence of non-collision singularities in the n -body problem. *Ann Math* 135(3):411–468
- Zizzi P (2004) Spacetime at the Planck scale: the quantum computer view. In: Garola C, Rossi A, Sozzo S (eds) *The foundations of quantum mechanics: historical analysis and open questions—Cesena 2004*, pp 345–358. World Scientific, Singapore. Preprint available Online, arXiv:gr-qc/0304032. Accessed 4 July 2008
- Zuse K (1967) Rechnender Raum. *Elektronische Datenverarbeitung* 8:336–344. Available Online, <ftp://ftp.idsia.ch/pub/juergen/zuse67scan.pdf>. Accessed 4 July 2008. This is a preliminary summary of Zuse (1969)
- Zuse K (1969) *Rechnender Raum*. Schriften zur Datenverarbeitung, Band 1. Friedrich Vieweg & Sohn, Braunschweig. English translation: *Calculating Space*, MIT Technical Translation AZT-70-164-GEMIT, MIT (Proj. MAC), Cambridge, MA. 02139, Feb. 1970. Available Online, <ftp://ftp.idsia.ch/pub/juergen/zuserechnenderraum.pdf>. Accessed 4 July 2008. A German-language summary appeared in 1967 as Zuse (1967)