1

Introduction

1.1 What is time?

This book explores the hypothesis that time is discrete rather than continuous. Time is an enigma, so we should expect some metaphysics and philosophy to creep into the discussion. Our inclination is to avoid those disciplines as much as possible, so let us deal with them right now.

Metaphysics and philosophy deal with statements and conjectures that cannot be empirically validated. In those disciplines there are constant references to absolutes such as existence, good and bad, and suchlike without further qualification, as if everyone accepted them as meaningful concepts. Absolutes are the key things we wish to avoid. For the record, we define an absolute statement as one that is considered to be true regardless of any caveats or criteria, i.e., context-free. In contrast, a contextual statement has a truth value that is meaningful only relative to its particular context.

The idea that physical truth can be contextual is an unfamiliar and uncomfortable one to physicists conditioned to believe that the laws of physics transcend the context of observation because they can be empirically validated. In fact, that is a circular line of reasoning. Every experiment is defined by its own context and experimentalists have to work hard to create that context: the search for the Higgs particle at the Large Hadron Collider did not happen overnight. Because it is impossible to actually perform all imaginable experiments, the known laws of physics have been validated only relative to a finite subset of all possible contexts. Therefore, the laws of physics are contextual, not absolute. It is metaphysics to think otherwise.

Despite that, there are numerous examples in physics of a conditioned metaphysical belief in an absolute. Lorentz covariance in special relativity (SR) is the principle that the laws of physics apart from gravitation take the same form in every inertial frame. In general relativity (GR) the corresponding concept is encoded into general covariance, the principle that the laws of physics are invariant with respect to arbitrary coordinate transformations. Are these absolute principles? Other examples come to mind: in thermodynamics, physicists make frequent references to the 'absolute temperature' of a system under observation (SUO), whilst in quantum mechanics (QM) they refer to 'the' probability of a quantum outcome. In fact, the temperature of an SUO is contextual on that SUO being in thermal equilibrium, whilst a probability in QM is always a conditional probability, i.e., contextual. As for Lorentz covariance and general covariance, these are more and more frequently these days being seen by cosmologists as useful guidelines in the construction of Lagrangians rather than absolute principles.

This issue impinges on us here because a potential criticism of discrete time (DT) mechanics is that it breaks Lorentz symmetry explicitly. That is true: we need to choose a preferred inertial frame in which to discretize time. We are not unduly concerned by this criticism, however, for several reasons. Three of these are as follows: (i) there is no empirical proof that time is continuous or otherwise; (ii) the aforementioned criticism does not take into account the empirical fact that we **can** use the laws of physics to identify a preferred local inertial frame anywhere in the Universe, the local frame relative to which the dipole anisotropy of the cosmic background radiation field vanishes (Cornell, 1989);[1] and (iii) conventional theories that are based on Lorentz symmetry are riddled with mathematical divergences, and DT may be a possible technique to grapple with them.

With the above in mind, we shall take as a guiding principle the view that there are no absolutes in physics: every concept or statement in physics should be accompanied by a statement of the context relative to which that statement's truth value makes sense.[2] Care should be taken to understand the opposite of relative truth: if a statement is not true relative to a given context, then it is false only relative to that context. Outside of that context, we should say *nothing*.

When we discuss any theory, the above principle of contextuality requires us to clarify the context in which our theory is to be discussed and held to be meaningful. In the case of DT, we should establish (i) *who* or *what* sort of observer is formulating the theory, (ii) for what purpose and to which ends the theory has been constructed, (iii) the principles underpinning the theory, including its limitations, and (iv) what might be done with the theory. We address these points in turn.

Point (i) In this book, time is studied from the perspective of the mathematical physicist, with no hidden agenda or philosophy. The reader will not

[1] This is a debatable point, in that it could be argued that the appearance of such a frame is a consequence of the laws of physics, not a fundamental feature in itself. But we would argue that the assertion that this frame was chosen at random by a quantum fluctuation is itself a metaphysical statement.

[2] Contextuality applies to mathematics as much as it does to physics. If it did not, why then do mathematicians spend time defining axioms and postulates? Theorems are true only relative to the relevant mathematical context.

be asked to believe either in continuous time (CT) or in DT. Since most of science deals with CT theories, it seems reasonable to redress the balance by investigating the consequences of DT mechanics. At our disposal will be mathematics backed up by intuition and supported by some empirical knowledge about time, such as its ordering property, its irreversibility and SR time dilation.

Point (*ii*) Many questions remain unanswered about the physical Universe, particularly the nature of space and time. We are not even sure how to classify time. Is it an object or a process? On the one hand, empty spacetime appears to have intrinsic physical properties such as curvature and vacuum polarizability, with particles being no more than quantum excitations of a basic state of empty space known as the vacuum. On the other hand, Mach's principle (Mach, 1912) and recent interpretations of QM (Rovelli, 1996) propose that space and time should be discussed in relational (contextual) terms. Which view is correct?

Even when spacetime is considered to be more than a relationship between objects, its structure remains debatable. Newtonian mechanics models space as a three-dimensional Euclidean manifold and time as a real line, whereas SR and GR merge space and time into a four-dimensional continuum known as spacetime. Although Einstein did acknowledge a debt to Mach's principle (Einstein, 1913), it is clear that GR spacetimes have intrinsic properties that can be measured, such as curvature. In GR, time is often identified with one of the four possible coordinates in a chosen coordinate patch and is continuous in that context.

On the other hand, some models of spacetime, such as Snyder's quantized spacetime (Snyder, 1947a, 1947b), suggest that continuous spacetime models may be too simplistic. Snyder's work motivated our particular interest in DT as an alternative to CT.

Point (*iii*) The principles we shall use are not controversial, apart from the single step of replacing the temporal continuum with a discrete set. All the standard principles of classical mechanics and QM adapted to DT are used in this book.

Point (*iv*) As for what DT can do for us, that remains to be seen. There are some nice things it can do for us, such as provide a natural (to the theory) cutoff in particle momentum. This may help cure some of the problems in CT quantum field theory, where the lack of any bound to linear momentum leads to divergences in loop integrals. This will be discussed in this book.

1.2 The architecture of time

We come now to a question central to this book: what is the architecture or structure of time? What sort of mathematical model best fits our intuitive notion of time?

This model or architecture should mirror the view of what we believe time represents and should incorporate into its rules whatever properties we believe time has. The architecture of time depends, therefore, on our beliefs about the Universe and how it runs. For instance, we might not believe that there is a single continuous strand of time such as Newton's absolute time (Newton, 1687). We might think there are many parallel strands of time each associated with a particular observer. Modelling such a 'multi-fingered architecture' would require the mathematics of parallel computer processing rather than the mathematics associated with single-processor computers.

In the following subsections we review several of the properties that the time concept should incorporate.

1.2.1 Events

Whatever model we decide on, it is a safe bet that we will incorporate into it the concept of an **event**.

Definition 1.1 An event is a well-defined, localized region of time and space, relative to a given observer.

Without the concept of an event, it would be impossible to discuss atoms and molecules, for instance. In general, events are assigned specific times and locations relative to a given observer. The existence of events is a supposition predicated on our world view. Whilst some quantum theorists view the Universe holistically as an enormous entangled state, quantum separability seems essential (Eakins and Jaroszkiewicz, 2003).

In particular, we should be aware of any hidden assumptions that we might be making about the nature of physical reality, as classically conditioned theorists find to their cost when they try to explain experiments such as the famous double-slit experiment (Tonomura and Ezawa, 1989). The quantum explanation of this experiment is at odds with the metaphysical classical explanation that an electron impacting on the final detecting screen had taken one or other of two possible paths on its journey from the source to that detecting screen. Quantum mechanically, we are not entitled to hold such a view. Therefore the following question arises: has a definite path been taken or not?

Classically we would have to believe that it had, because in our mind's eye we imagine a classical particle always follows a unique, continuous trajectory from source to screen. Quantum principles, however, require us to say nothing about this, *if* we have not attempted to detect anything about which path was taken.

For reasons such as this we have included the reference to an observer in Definition 1.1, for, if we did not, we would be implying that events could exist and things could happen regardless of who or what was observing them. That is a very classical, absolutist perspective on reality.

1.2.2 Temporal ordering

Suppose now that an observer had detected two or more events. How would that observation be modelled mathematically? In addition to any other attributes such as position, the observer would assign a time to each event, that is, a real number, which we call the assigned time (relative to that observer). Now, a crucial property of real numbers is that they are an ordered set. If we pick any two real numbers x and y, then only one of the following three statements can be true: x is less than y, or x equals y, or x is greater than y.

This means that there is temporal ordering relating any two observed events. If event A is assigned a time t_{A} and event B is assigned a time t_{B}, then $t_{\mathrm{A}} < t_{\mathrm{B}}$, or $t_{\mathrm{A}} = t_{\mathrm{B}}$, or $t_{\mathrm{A}} > t_{\mathrm{B}}$. These mathematical statements are interpreted physically as follows. In case (i) we say that A is earlier than B (or, equivalently, that B is later than A), whilst in case (ii) we say that A and B are simultaneous.

When SR emerged into the general consciousness of physicists, a significant conceptual problem for theorists conditioned to believe in Newtonian absolute time (Newton, 1687) was that simultaneity in SR is contextual. In contrast to Newtonian mechanics, where all classical observers agree on the relative temporal ordering of all events, SR asserts that A could be earlier than B relative to one observer and later than B relative to another.

The loss of absolute simultaneity in SR concerns two or more observers. We may bypass this issue by the simple method of restricting our attention to a single observer. In that context, all observed events have a well-defined temporal ordering relative to that observer.

1.2.3 Causality

A new factor now enters into the discussion: cause and effect. Suppose we have two events A and B, with A earlier than B according to some observer. That observer may have reason to believe in a causal link between A and B, in that there may be evidence in support of the notion that A caused B, or at least had some influence on B.

The notion of causality is notoriously difficult to pin down, principally because it requires us to contemplate counterfactuality, that is, valid logical conclusions that are based on premises known to be false. The 'mark-method' of Reichenbach (1958) demonstrates the point clearly (Whitrow, 1980). Reichenbach considers two events A and B, with A regarded as the cause of B. This relationship is denoted by AB, with the left–right ordering implying causal association. Now suppose that what happened at A had been slightly altered. We indicate this by marking the symbol A with an asterisk, i.e., A is replaced by A^*. Then one of two things could happen: either B is unchanged or else B is changed to B^*. Reichenbach asserts that the combinations AB, AB^* and $\mathrm{A}^*\mathrm{B}^*$ are consistent with A being the cause of B, but $\mathrm{A}^*\mathrm{B}$ is inconsistent with A being the cause of B.

The problem with this line of reasoning is that it is based on classical counterfactuality, which assumes counterfactual arguments are empirically meaningful. This is not the case in QM. A much quoted dictum attributed to John A. Wheeler states the quantum position elegantly: ‘No elementary phenomenon is a phenomenon until it is an observed (or registered) phenomenon’ (Wheeler, 1979).

1.2.4 The dimensions of time

Time is generally regarded as having a single dimension, but the question of physics based on two or more times has been discussed by experimentalists and theorists. We give an example of each.

Several decades ago, the astrophysicist Tifft measured the red shift of distant galaxies and came to the conclusion that redshifts were ‘quantized’; that is, they appeared to be clustered into groups or bands. Subsequently, he developed an interpretation of his data that was based on a model of three-dimensional time. In his model, he asserted that ‘each galaxy evolves along a 1-d timeline such that within a given standard galaxy standard 4-d space-physics is satisfied. The model deviates from ordinary physics by associating different galaxies with independent timelines within a general 3-d temporal space.’ (Tifft, 1996). In the model, temporal quantization, involving photon exchange between galaxies and observers, was invoked to account for the discrete structures in his redshift data.

It would be unfair to criticize this approach since it is no more than an attempt to fit an unusual mathematical model to actual observations. Unfortunately, although Tifft’s data were consistent with some subsequent observations, the most recent analysis concludes that there is no periodic structure in the redshift data (Schneider *et al.*, 2007). Therefore, the idea that time may be part of a three-dimensional continuum appears incorrect.

Tifft’s model incorporates a serial time of the form that we are used to, since a worldline in any dimensional spacetime can be parametrized by a single real variable, which can be called a time.

We come now to a theoretical discussion by the theorist Tegmark of the mathematical consequences of having a genuine multi-dimensional form of time.

Tegmark analyses a flat spacetime with p time dimensions modelled by coordinates $\mathbf{t} \equiv \{t^1, t^2, \ldots, t^p\}$ and q spatial dimensions modelled by coordinates $\mathbf{x} \equiv \{x^1, x^2, \ldots, x^q\}$ (Tegmark, 1997). The important property here is the signature of the metric, denoted by (p, q). Drawing on experience in standard-signature $(1, 3)$ SR spacetime, or Minkowski spacetime, Tegmark’s discussion focuses on a second-order partial differential wave equation for a spinless relativistic field φ of the form

$$\left\{ \sum_{i=1}^{p} \frac{\partial^2}{\partial t^i \, \partial t^i} - \sum_{j=1}^{q} \frac{\partial^2}{\partial x^j \, \partial x^j} \right\} \varphi(\mathbf{t}, \mathbf{x}) + V(\varphi(\mathbf{t}, \mathbf{x})) = 0, \tag{1.1}$$

where V is some self-interaction term, such as a mass term, which does not depend on any derivatives of the field φ. The point here is the sign difference between the timelike coordinates $\{t^1, t^2, \ldots, t^p\}$ and the spacelike coordinates $\{x^1, x^2, \ldots, x^q\}$ in (1.1), which arises from an assumed line element of the form

$$ds^2 = (dt^1)^2 + \cdots + (dt^p)^2 - (dx^1)^2 - \cdots - (dx^q)^2. \tag{1.2}$$

The sign change in (1.2) is of the greatest importance to the modelling and interpretation of physics. In the case $p = 0$, referred to as the elliptic case, observers have no predictive power (Tegmark, 1997). There are no lightcones and no timelike worldlines in such spaces. This case corresponds to the imaginary-time scenario discussed by Minkowski in 1908 (Minkowski, 1908), which is frequently invoked in various branches of cosmology and particle physics in attempts to regularize mathematical divergences. There are numerous issues about this scenario that should cause concern (Jaroszkiewicz, 2002).

In the case of our Universe as we believe it to be, $p = 1$ and $q = 3$. Then the above differential equation is an example of a hyperbolic differential equation (Arfken, 1985). This case models the physically reasonable situation where an observer can use initial data over an initial spacelike hypersurface in relativistic spacetime to predict the final data over a final spacelike hypersurface. There are lightcones and timelike worldlines in such a spacetime. Tegmark concludes that the case $q < 3$ gives too simple a model and the case $q > 3$ leads to instability in the physics.

The remaining possibility, $p > 1$, is known as the ultrahyperbolic regime and leads to unpredictability.

Tegmark's analysis is based on what observers might see or be unable to see for various values of p and q, the value $p = 1$ being consistent with information flow in the form we are used to. In other words, observational criteria are used to decide what the spacetime architecture of the Universe might be.

There is no principle in GR that forbids a change of signature, apart from considerations such as those of Tegmark. The possibility that the signature changes dynamically has been considered. For instance, particle production from signature change from $(1, 1)$ to $(0, 2)$ was discussed by Dray *et al.* (1991).

1.2.5 Manifold time versus process time

The question raised earlier, namely that of whether time is an object or a process, leads to two mutually exclusive interpretations of time referred to as manifold time and process time, respectively (*Encyclopædia Britannica*, 2000). Manifold time regards time as a geometrical quantity, an objective thing having a single dimension and all the ordering properties of the real line. Manifold time represents an absolutist approach to time. An associated concept is the block universe (Price, 1997), which models spacetime as an object.

On the other hand, process time models time contextually: time is not anything that exists by itself but is an attribute of physical processes, a manifestation of change. Since change can be defined only relative to the memories of observers or their equivalent, process time implies the presence of observers and is compatible with and consistent with relational QM (Rovelli, 1996).

The difference between the manifold time and process time perspectives is important to us in this book. If time is indeed best described as part of a monolithic four-dimensional continuum, then discretizing time requires us to choose a preferred frame of reference.

On the other hand, if time is a manifestation or résumé of what a given observer experiences, then discretization of time need be considered only for *that* observer. An analogy can be drawn here with electron spin. In classical mechanics (CM), angular momentum is a continuous variable and a spinning particle can have any value of angular momentum. But any given observer detecting for quantized electron spin in a Stern–Gerlach experiment (Gerlach and Stern, 1922a, 1922b) can assign to it only one of two possible quantum spin values, a discretization of continuous angular momentum sometimes referred to as spatial quantization. The classical spatial continuum still exists in the formalism because the orientation parameters of the apparatus are not quantized, i.e., the direction in space of the main magnetic field axis is classical and can take on any value in QM.

The history of physics contains two important examples analogous to the manifold–process time debate: (i) classical thermodynamics treats temperature and entropy as classical attributes of continuous matter, whereas statistical physics interprets both of these concepts as statistical attributes of ensembles of systems in thermal equilibrium; and (ii) heat is interpreted as a substance in the theory of phlogiston, whereas the modern view is to regard it as part of a process.

An example of such an idealogical conflict in mathematics comes from probability theory, where the frequentist view of probability as an absolute quality of a random variable that can be measured approximately by sampling contrasts with the Bayesian view of probability as conditioned by prior information, i.e., a contextual approach to probability.

1.2.6 Multi-fingered time

Once we think of time as a manifestation of processes involving observers, we are naturally led to the idea that there may be as many times as there are observers. In SR, this is a well-understood feature of proper time: different observers following different worldlines experience time in a local, path-dependent way. This is the source of the so-called twin paradox, which is a paradox only if time is interpreted in the wrong way.

The multi-fingered time interpretation is compatible with relational QM (Rovelli, 1996). Moreover, time as it relates to the process of observation is naturally discrete: an observer prepares a quantum state by an initial time and registers outcome information at a subsequent final time. These times are reasonably well defined relative to that observer.

1.2.7 Temporal continuity

In addition to the above interpretational issues, there arises the question of which specific mathematical structures should be used in the modelling of time. Immediately we are faced with two choices: is time continuous or discrete? In CT, time is represented by a continuous parameter usually ranging over some interval $I_{\mathrm{if}} \equiv [t_{\mathrm{i}}, t_{\mathrm{f}}]$ of the real line, where t_{i} is the initial time and $t_{\mathrm{f}} > t_{\mathrm{i}}$ is the final time of some experiment. The universal convention in physics, which we follow in this book, is that, given two different values t_{i} and t_{f} of time, the larger value represents a physically later time in the laboratory, with 'later' being associated with the observed direction of the expansion and evolution of the Universe. On the other hand, DT is modelled as a sequence $\{t_n\}$ of real numbers, labelled by an integer n running from an initial value M to some final value $N > M$.

CT remains a powerful and popular model, which from the time of Newton onwards has been thoroughly explored and exploited. On the other hand, DT is still being developed. All the indications are that the importance of DT is growing, particularly on account of the impossibility of modelling CT exactly on a computer. Many CT models are approximated by appropriate discretizations of time so that they can be modelled on computers.

Our aim in this book is to discuss DT in those areas with which we are most familiar, but the importance of CT should not be overlooked. CT will be a central feature in much of our discussion and frequently used alongside DT as a parallel component of the discussion. It is possible to discuss the two views of time in the same context, provided that care is taken. To illustrate what we mean, consider a stone skipping over the surface of a pond. The pond's surface can be regarded as a continuum, but where the stone bounces off that surface is described as a discrete set.

Because CT is a central element of Newtonian CM, we may reasonably assume it is familiar to the reader. However, we shall review some of its basic features in order to highlight the differences between it and DT.

To understand CT we should understand the definition of a linear continuum. A continuum is a space, i.e. a set with certain properties. We need not concern ourselves with the nature of the points of the space.

Definition 1.2 A partially ordered set, or poset, S, is a set with a binary relation denoted by $\leqslant$, such that we have

(i) **reflexivity**: for every element x in S, $x \leqslant x$;
(ii) **antisymmetry**: if x and y are elements of S such that $x \leqslant y$ and $y \leqslant x$, then $x = y$; and

(iii) **transitivity**: if x, y and z are elements of S such that $x \leqslant y$ and $y \leqslant z$, then $x \leqslant z$.

Posets are important in SR because of the following exotic possibility: there may be elements u, v for which the binary relation $\leqslant$ is not defined. Minkowski spacetime $\mathcal{M}^4$, the four-dimensional spacetime of SR, has a Lorentzian metrical lightcone structure that creates this possibility. We can pick pairs of different events U, V in $\mathcal{M}^4$ such that, in some inertial frames, the times t_U, t_V assigned to them respectively satisfy $t_U \leqslant t_V$, and such that, in other inertial frames, we have $t'_V \leqslant t'_U$. Such pairs of events will be called relatively spacelike pairs.

We shall return to posets in Chapter 29, when we discuss causal sets.

To define a linear continuum, we need extra conditions. In particular, we need to eliminate the possibility of relatively spacelike relationships by introducing an extra condition, which turns a poset into a totally ordered set.

Definition 1.3 A **linearly ordered** or **totally ordered set** S is a poset with the additional property of

(iv) **totality**: for any two elements x, y of S, then $x \leqslant y$ or $y \leqslant x$.

The totality property of a totally ordered set essentially places a veto on finding a spacelike pair in S. An additional binary relation can be defined for each totally ordered set.

Definition 1.4 A **strict total order** on a totally ordered set is a binary relation $<$ such that $x < y$ if and only if $x \leqslant y$ and $x \neq y$.

Before we define a linear continuum, we need to define the concept of least upper bound.

Definition 1.5 Let S be a subset of a poset X. Then an element b of X is an **upper bound** for S if, for every element x of S, $x \leqslant b$.

Definition 1.6 Let S be a subset of a poset X. If there exists an element b_0 of X such that $b_0 \leqslant b$ for every upper bound of S, then b_0 is the **least upper bound** or **supremum** for S.

We note that the supremum, if it exists, is unique.

We now have the structures needed to define a linear continuum.

Definition 1.7 A *linear continuum* is a non-empty totally ordered set S such that

(1) S has the least upper bound property; and
(2) given any two different elements x and y of S such that $x < y$, there always exists another, distinct element z in S such that $x < z$ and $z < y$ (we write $x < z < y$).

Property (2) is at the heart of the difference between CT and DT. Suppose we have two values of time, t_1 and t_2, such that $t_1 < t_2$. If we know we are