Foundations of Physics, Vol. 36, No. 9, September 2006 (© 2006)
DOI: 10.1007/s10701-006-9064-x

A Nontemporal Probabilistic Approach to Special and General Relativity

Frank Blume[1]

Received June 5, 2005; revised April 10, 2006; Published online August 12, 2006

We introduce a discrete probabilistic model of motion in special and general relativity that is shown to be compatible with the standard model in the statistical limit.

KEY WORDS: relativity; motion; probability; nontemporal.

1. INTRODUCTION

Anyone familiar with the fundamental principles of the differential calculus may have wondered at times whether the assumption of continuity in space and time, on which so many of our mathematical models are elegantly based, is indeed a true reflection of physical reality. Is it not conceivable that our common notion of temporal and spacial continuity is just as illusory as the impression of seamless change created by the rapid succession of discrete static images on a movie screen? The question is not new, and fairly detailed attempts to develop nontemporal discrete models of physical reality have been made (see, e.g., Ref. 1), but what we hope to offer in the present paper is a novel approach to the problem of discretizing special and general relativity that perhaps adds another facet to already existing interpretative schemes.

For the most part, our discussion will focus on specific mathematical derivations, but in order to provide the reader with an adequate conceptual framework, a few introductory remarks of a more general philosophical nature are no doubt appropriate.

To begin with, we will revisit the familiar twin "paradox" of special relativity: the proper times associated with two paths that connect the

[1] John Brown University, Siloam Springs, AR 72761, USA. e-mail: fblume@jbu.edu

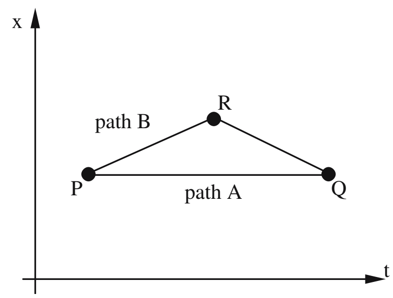

Fig. 1. Nonsymmetric paths in space–time.

same starting and endpoints P and Q in Minkowski space (see Fig. 1) are in general not equal. The mathematical explanation for this surprising fact is, of course, well known, and all protests of common sense are easily countered by pointing out that the human cognitive apparatus, formed under the essentially nonrelativistic conditions of earthly life, is most likely ill equipped to intuitively grasp the properties of space–time in its more extreme manifestations. It may therefore appear that the twin "paradox" is perfectly well understood—and in a sense it certainly is—but upon further reflection we quickly find ourselves confronted with some perplexing questions. Consider for instance the problem of how the difference in proper time along the paths A and B in Fig. 1 is encoded. If the paths are traced out by human observers, this difference may come to light in a simple comparison of clocks carried by these observers, but what if we examine instead two elementary particles, say, two electrons? Which of them is younger and which is older[2] by the time they arrive at Q? (Note: whether, in fact, quantum mechanics allows us to speak of an electron tracing out a particular path shall be of no concern to us for the moment.) More generally, we may ask how the passing of time manifests itself, not only with regard to phenomena involving relative motion, but universally, as the ordering principle underlying all reality.

To explore this issue, we first observe that any measurement of time, commonly performed, is indirect. What we record are certain material *changes* happening in time, such as the changes in the position of a hand on a stopwatch, rather than time itself. It is by no means obvious what kind of metaphysical essence we should hope to discover in a purely

[2] It must be emphasized that in raising the question of determining the "age of an electron" we are not trying in any way to put in doubt the rich empirical evidence in favor of relativistic time dilation that exists especially in the realm of elementary particles. Our intent is rather to probe the twin "paradox" at a somewhat deeper, more philosophical level.

abstract conception of time absent any perception of change. Is it meaningful to speak of the passing of time if nothing is happening and nothing is perceived? Faced with questions such as these that pertain to the ultimate nature of reality, we cannot expect to find final conclusive answers, but for the sake of the present argument, we will assume that the notions of time and change are indeed synonymous.

Having thus identified our point of departure, we now return to the image of two electrons following the world lines A and B, respectively, in the space–time diagram in Fig. 1. If, as predicted by special relativity, the proper time associated with A is greater than that associated with B, and if any progression in time requires the occurrence of some sort of change, then what exactly are the changes happening to the electron on path A that are not happening to the one on path B, or at least are not happening at the same rate? Since the diagram in Fig. 1 is of little use as we try to address this question, we boldly—and, perhaps, insanely—propose to replace the smooth world lines A and B by erratic random walks traced out at the speed of light (see Fig. 2 for a rather crude illustration). More specifically, we suggest that any uniform linear motion at velocity v is to be represented by a random walk which in turn is represented by a sequence of independent identically distributed (i.i.d.) binomial random variables $x_1, \ldots, x_n$ such that

$$P(x_i = 1) = p \quad \text{and} \quad P(x_i = -1) = 1 - p \quad \text{for all } i \in \{1, \ldots, n\}. \tag{1}$$

Given this setup, the average change in position over a time interval of length 1 (as measured in the inertial frame of reference in which the random walk is recorded) is the sample mean

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \tag{2}$$

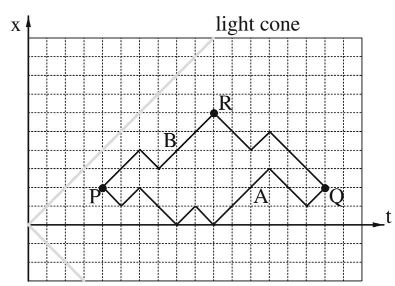

Fig. 2. Random walks corresponding to the paths A and B.

and the expected value of $\bar{x}$ or, equivalently, of any one of the random variables x_i is therefore the velocity v, i.e.,

$$v = E(\bar{x}) = E(x_i) = 2p - 1. \tag{3}$$

Consequently, the probability p is to be defined by the equation

$$p = \frac{1 + v}{2}. \tag{4}$$

The model that we are in the process of developing here is, of course, highly simplistic and not entirely convincing, but for the purpose of illustrating certain fundamental concepts that underlie also the more adequate mathematical constructions in Secs. 2 and 3 it actually is well suited. The idea, for example, that time is synonymous with change finds its natural expression in the identification of proper time with randomness. To establish this link we consider the sample variance

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{5}$$

which is commonly used as a measure of statistical spread. Elementary probability theory teaches us that the expectation of S^2 is equal to the variance σ^2 of the individual random variables x_i, which in our case is given by the equation

$$\sigma^2 = \mathrm{Var}(x_i) = E((x_i - E(x_i))^2) = E(x_i^2) - E(x_i)^2 = 1 - v^2. \tag{6}$$

Since the proper time associated with a uniform linear motion at velocity v over a time interval of length n is known to be

$$\tau = n\sqrt{1 - v^2}, \tag{7}$$

and since the term on the right is equal to

$$n\sigma = n\sqrt{\mathrm{Var}(x_i)}, \tag{8}$$

we may infer that a good approximate measure of proper time is the sample standard deviation S multiplied by n (at least in the special case of a uniform linear motion).

As far as our earlier problem of comparing the age of the two electrons upon arrival at Q is concerned, we can now offer an elegant solution. Since the velocities along the segments $\overline{PR}$ and $\overline{RQ}$ that make up path B are both different from zero, and since the velocity along path A is equal to zero, it follows that the value of the sample standard deviation S can be expected to be greater for A than for either one of the segments $\overline{PR}$ and $\overline{RQ}$. In other words, of the two given paths, A is most likely the one with a higher incidence of random changes in the direction of travel, and where there is more change, more time is passing. By implication, the electron on path A may indeed be expected to "grow older" more rapidly, and the difference in "age" is encoded in the structures of the corresponding random walks.

Having thus reduced time to a measure of randomness we may, if we wish, dispose of it altogether and adopt a genuinely nontemporal point of view. Such an interpretative shift appears plausible also in light of the fact that in our model random walks are always traced out at the speed of light (see Fig. 2) and that the standard computation of proper time (via an integral of $\sqrt{1 - v(t)^2}$) therefore yields the value zero. However, the question of whether time does or does not exist is by no means essential to the mathematical derivations that will be the focus of our discussion in Secs. 2 and 3. In the end it is up to the reader to decide which mode of interpretation—temporal or nontemporal—seems more plausible.

Another interesting aspect of our probabilistic scheme is the fact that the speed of light $c = 1$ is naturally found to be the largest attainable speed. This fact, which in standard treatments of special relativity is adopted as an axiom, is a direct consequence of Eq. (3) because the minimal and maximal probabilities $p = 0$ and $p = 1$ are here easily seen to correspond to the minimal and maximal velocities $v = -1$ and $v = 1$. As trivial as this observation may seem, it nevertheless highlights a crucial distinction between classical and relativistic kinematics in that it shows the feasibility of our probabilistic approach to be dependent upon the existence of a universal upper bound for $|v|$. We would be overstating our case if we now asserted that special relativity is intrinsically probabilistic, but to say that our model is well adjusted to the structure of special relativity seems fair.

Given the outline provided in the preceding paragraphs, we will now discuss some possible objections. First of all, there is the problem that special relativity does not allow for an object with a nonvanishing rest mass to travel at the speed of light. In particular, the requirement that all moves in a given random walk be performed at exactly the speed of light is apparently impossible to satisfy. To this we reply that the proposed probabilistic picture will be essentially indistinguishable from the standard continuous picture if only the unit step length $|x_i| = 1$ is sufficiently small (the Planck length comes here to mind). In other words, we simply evade the question by demanding a level of resolution that is beyond the reach of empirical verification. To readers who fail to be impressed by this retreat into an inaccessible microrealm we would like to point out that our intent is not to genuinely revise the theory of relativity or to formulate new experimentally verifiable hypotheses, but only to suggest a few alternative interpretations.

A more serious objection concerns the transformation of random walks as we pass from one inertial frame of reference to another. When we introduced the idea that uniform linear motions are to be represented by sequences of random variables x_i with $|x_i| = 1$ there was no mention of transformations to different coordinate frames at all. This suggests that the proposed representation was intended to be universally valid. However, distances in special relativity are not absolute and, in particular, measurements of step lengths in a given random walk cannot, so it seems, yield the value $|x_i| = 1$ independent of an observer's state of motion. Our response in this case is of central importance for everything that follows:

> The very nature of our probabilistic model is such that a a direct spatiotemporal correspondence between random walks is given up in favor of a mere statistical correlation. (9)

In other words, an observer's experience of the world is essentially subjective, and a coherent picture emerges only at the level of statistical averages. Perhaps the best way to interpret this principle is by way of analogy to quantum mechanics where probabilities are known, but individual paths are essentially ficticious. So we may want to assume that observers can predict approximate positions but are not able to trace each individual step in a given random walk.

Keeping this fundamental idea in mind, we are now ready to engage in a more serious discussion of the mathematical structure of a probabilistic theory of relativity.

2. PROBABILITY DENSITIES IN SPECIAL RELATIVITY

The main reason why the probabilistic model developed in the Introduction was so utterly simple is that in one spatial dimension there are only two possible directions of motion—the positive and the negative. By contrast, the number of available directions in two- or three-dimensional space is infinite. In order to describe motion in a higher-dimensional setting, it will therefore be necessary to work with continuous rather than discrete random variables.

Considering first the case of two spatial dimensions, we assume that we are given a probability density function p on the unit circle

$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \tag{10}$$

such that the probability for a move in a random walk to fall within an angular range from θ to $\theta + d\theta$ is

$$p(\mathbf{h}(\theta))d\theta \tag{11}$$

where

$$\mathbf{h}(\theta) = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \tag{12}$$

for all $\theta \in [-\pi, \pi]$.

To simulate a uniform linear motion in the xy-plane we further assume that $\theta_1, \ldots, \theta_n$ is a sequence of i.i.d. random variables such that for each index i the density of θ_i is $p(\mathbf{h}(\theta))$. Setting

$$\begin{aligned} x_i &:= \cos(\theta_i), \\ y_i &:= \sin(\theta_i), \end{aligned} \tag{13}$$

the two-dimensional random walk that we associate with the motion in question is described by the displacement vectors

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \mathbf{h}(\theta_i). \tag{14}$$

Since each move in a random walk is supposed to be performed at the speed of light, and since each displacement vector has length 1, we may infer that the expected velocity is

$$\mathbf{v} := \begin{pmatrix} E(x_i) \\ E(y_i) \end{pmatrix} = \begin{pmatrix} E(\cos(\theta_i)) \\ E(\sin(\theta_i)) \end{pmatrix} \text{(independent of } i\text{)} \tag{15}$$

where

$$\begin{aligned} E(\cos(\theta_i)) &= \int_{-\pi}^{\pi} \cos(\theta)\, p(\mathbf{h}(\theta))\, d\theta, \\ E(\sin(\theta_i)) &= \int_{-\pi}^{\pi} \sin(\theta)\, p(\mathbf{h}(\theta))\, d\theta. \end{aligned} \tag{16}$$

Furthermore, just as in the one-dimensional case, the proper time corresponding to a particular move is again found to be a natural measure of randomness because

$$\begin{aligned}\sqrt{1-\|\mathbf{v}\|^2} &= \sqrt{1-(E(x_i)^2+E(y_i)^2)} \\ &= \sqrt{E(x_i^2+y_i^2)-(E(x_i)^2+E(y_i)^2)} \\ &= \sqrt{\mathrm{Var}(x_i)+\mathrm{Var}(y_i)}\end{aligned} \tag{17}$$

for all $i \in \{1, \ldots, n\}$.

Since the probabilistic model outlined above is still too general to be of any use, it is advisable that we now discuss transformations of density functions as we pass from one inertial frame of reference to another. So suppose that two observers, A and B, are moving relative to each other at the constant velocity $\mathbf{v}$, as measured by A. The space–time coordinates of the inertial systems associated with these two observers are denoted by $t,\ x,\ y$ and $\tau,\ \xi,\ \eta$, respectively. Given this setup, our objective is to determine how a probability density q on the unit circle $\{(\xi, \eta) \mid \xi^2+\eta^2 = 1\}$ that describes the random walk of an object C, as viewed by B, would have to be transformed in order to describe the corresponding random walk in the inertial system of A. Considering the very general case where two coordinate frames are rotated but not shifted relative to each other (see Fig. 3), we may infer the existence of a 2×2 orthogonal matrix $\mathbf{A}$ (i.e., $\mathbf{A}^t\mathbf{A} = \mathbf{I}$) such that the Lorentz transformation is described by the equations

$$\begin{aligned}\tau &= \gamma(t-\mathbf{r}\cdot\mathbf{v}), \\ \begin{pmatrix}\xi \\ \eta\end{pmatrix} &= \mathbf{A}(\mathbf{r}+(\gamma-1)\mathbf{P}_{\mathbf{v}}(\mathbf{r})-\gamma t\mathbf{v})\end{aligned} \tag{18}$$

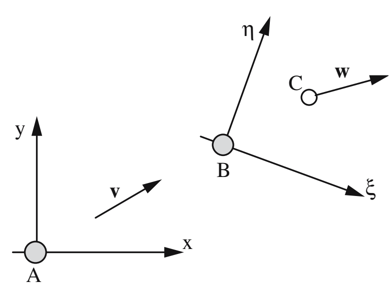

Fig. 3. Relative motion of two inertial observers.

where

$$\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \gamma = \frac{1}{\sqrt{1 - \|\mathbf{v}\|^2}}, \quad \text{and} \quad \mathbf{P}_{\mathbf{v}}(\mathbf{r}) = \frac{(\mathbf{r} \cdot \mathbf{v})\mathbf{v}}{\|\mathbf{v}\|^2}. \tag{19}$$

Furthermore, if A observes a spatial displacement by a vector $\mathbf{r}$ at the speed of light, then the corresponding time coordinate is $t = \|\mathbf{r}\|$ and, according to (18), the spatial displacement observed by B (which, of course, also occurs at the speed of light) is

$$\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \mathbf{A}(\mathbf{r} + (\gamma - 1)\mathbf{P}_{\mathbf{v}}(\mathbf{r}) - \gamma \|\mathbf{r}\| \mathbf{v}) =: \mathbf{g}(\mathbf{r}). \tag{20}$$

Note: unlike the Lorentz transformation, $\mathbf{g}$ is not linear, but it does satisfy the equation $\mathbf{g}(\lambda \mathbf{r}) = \lambda \mathbf{g}(\mathbf{r})$ for all $\lambda \geqslant 0$.

With the introduction of $\mathbf{g}$ we actually are going to restore the sort of direct spatiotemporal correspondence that we claimed to have given up with the principle stated in (9), but the reader may rest assured that the restoration is only temporary (despite the fact that time does not exist, but that's a different matter) and that in the end our precious principle will not be violated.

In order to determine the probability $p(\mathbf{h}(\theta))d\theta$ associated with a move in the random walk of C, as observed by A, we notice that the angular width $d\theta$, when viewed by B, assumes the value

$$\left\| \frac{\mathbf{g}(\mathbf{h}(\theta + d\theta))}{\|\mathbf{g}(\mathbf{h}(\theta + d\theta))\|} - \frac{\mathbf{g}(\mathbf{h}(\theta))}{\|\mathbf{g}(\mathbf{h}(\theta))\|} \right\| = \left\| \frac{d}{d\theta} \frac{\mathbf{g}(\mathbf{h}(\theta))}{\|\mathbf{g}(\mathbf{h}(\theta))\|} \right\| d\theta. \tag{21}$$

Thus, it appears that $p(\mathbf{h}(\theta))d\theta$ ought to be equal to, or at least proportional to

$$q\left(\frac{\mathbf{g}(\mathbf{h}(\theta))}{\|\mathbf{g}(\mathbf{h}(\theta))\|} \right) \left\| \frac{d}{d\theta} \frac{\mathbf{g}(\mathbf{h}(\theta))}{\|\mathbf{g}(\mathbf{h}(\theta))\|} \right\| d\theta. \tag{22}$$

However, we need to be careful, because in the inertial frame of B the probability above is associated with a random displacement of length $\|\mathbf{g}(\mathbf{h}(\theta))\|$, which in general is different from 1. To compensate for this relativistic spatial distortion and to thereby reinstate the principle in (9), we must divide by $\|\mathbf{g}(\mathbf{h}(\theta))\|$ the term in (22). Consequently, in normalizing the function

$$f(\theta) := \frac{q\left(\frac{\mathbf{g}(\mathbf{h}(\theta))}{\|\mathbf{g}(\mathbf{h}(\theta))\|} \right) \left\| \frac{d}{d\theta} \frac{\mathbf{g}(\mathbf{h}(\theta))}{\|\mathbf{g}(\mathbf{h}(\theta))\|} \right\|}{\|\mathbf{g}(\mathbf{h}(\theta))\|} \tag{23}$$

to unity (unit probability, that is) we may assert the density p to be given by the equation

$$p(\mathbf{h}(\theta)) = \frac{f(\theta)}{\int_{-\pi}^{\pi} f(\theta)\, d\theta}. \tag{24}$$

Since the definition of $\mathbf{g}$ in (20) readily implies that

$$\|\mathbf{g}(\mathbf{r})\| = \gamma(\|\mathbf{r}\| - \mathbf{r}\cdot\mathbf{v}), \tag{25}$$

we may infer that the derivative of $\mathbf{F} := \mathbf{g}/\|\mathbf{g}\|$ is

$$\begin{aligned} D\mathbf{F}(\mathbf{r}) &= \frac{D\mathbf{g}(\mathbf{r})}{\|\mathbf{g}(\mathbf{r})\|} - \frac{\mathbf{g}(\mathbf{r})\left(\frac{\mathbf{r}^t}{\|\mathbf{r}\|} - \mathbf{v}^t\right)}{\gamma(\|\mathbf{r}\| - \mathbf{r}\cdot\mathbf{v})^2} \\ &= \frac{\mathbf{A}\left(\mathbf{I} + \frac{(\gamma-1)\mathbf{v}\mathbf{v}^t}{\|\mathbf{v}\|^2} - \frac{\gamma\mathbf{v}\mathbf{r}^t}{\|\mathbf{r}\|}\right)}{\gamma(\|\mathbf{r}\| - \mathbf{r}\cdot\mathbf{v})} - \frac{\mathbf{g}(\mathbf{r})\left(\frac{\mathbf{r}^t}{\|\mathbf{r}\|} - \mathbf{v}^t\right)}{\gamma(\|\mathbf{r}\| - \mathbf{r}\cdot\mathbf{v})^2} \\ &= \frac{\mathbf{A}}{\gamma(\|\mathbf{r}\| - \mathbf{r}\cdot\mathbf{v})} + \frac{\mathbf{A}((\mathbf{r}\cdot\mathbf{v})\mathbf{v}\mathbf{r}^t - \|\mathbf{r}\|^2\mathbf{v}\mathbf{v}^t)}{(\gamma+1)\|\mathbf{r}\|(\|\mathbf{r}\| - \mathbf{r}\cdot\mathbf{v})^2} + \frac{\mathbf{A}(\|\mathbf{r}\|\mathbf{r}\mathbf{v}^t - \mathbf{r}\mathbf{r}^t)}{\gamma\|\mathbf{r}\|(\|\mathbf{r}\| - \mathbf{r}\cdot\mathbf{v})^2}. \end{aligned} \tag{26}$$

Since $\|\mathbf{h}(\theta)\| = 1$ and $\mathbf{h}(\theta)^t\mathbf{h}'(\theta) = 0$, it follows that

$$\begin{aligned} \frac{d}{d\theta}\frac{\mathbf{g}(\mathbf{h}(\theta))}{\|\mathbf{g}(\mathbf{h}(\theta)))\|} &= D\mathbf{F}(\mathbf{h}(\theta))\mathbf{h}'(\theta) \\ &= \frac{\mathbf{A}\mathbf{h}'(\theta)}{\gamma(1 - \mathbf{h}(\theta)\cdot\mathbf{v})} + \frac{\mathbf{h}'(\theta)\cdot\mathbf{v}}{(1 - \mathbf{h}(\theta)\cdot\mathbf{v})^2}\left(\frac{\mathbf{A}\mathbf{h}(\theta)}{\gamma} - \frac{\mathbf{A}\mathbf{v}}{\gamma+1}\right) \end{aligned} \tag{27}$$

and, in taking the square root of the dot product of this vector with itself, we obtain

$$\begin{aligned} \left\|\frac{d}{d\theta}\frac{\mathbf{g}(\mathbf{h}(\theta))}{\|\mathbf{g}(\mathbf{h}(\theta)))\|}\right\| &= \|D\mathbf{F}(\mathbf{h}(\theta))\mathbf{h}'(\theta)\| \\ &= \frac{\|\mathbf{h}'(\theta)\|}{\gamma(1 - \mathbf{h}(\theta)\cdot\mathbf{v})} = \frac{1}{\gamma(1 - \mathbf{h}(\theta)\cdot\mathbf{v})}. \end{aligned} \tag{28}$$

Combining this result with (23) and (25), we may conclude that

$$f(\theta) = \frac{q(\mathbf{F}(\mathbf{h}(\theta)))}{\gamma^2(1 - \mathbf{h}(\theta)\cdot\mathbf{v})^2}. \tag{29}$$

In order to evaluate the integral of f from $-\pi$ to π, we introduce the substitution

$$\mathbf{h}(\psi) = \mathbf{F}(\mathbf{h}(\theta)). \tag{30}$$

This yields

$$\left|\frac{d\theta}{d\psi}\right| = \frac{\|\mathbf{h}'(\psi)\|}{\|D\mathbf{F}(\mathbf{h}(\theta))\mathbf{h}'(\theta)\|} = \frac{1}{\|D\mathbf{F}(\mathbf{h}(\theta))\mathbf{h}'(\theta)\|}, \tag{31}$$

and therefore,

$$\begin{aligned}\int_{-\pi}^{\pi} f(\theta)\, d\theta &= \int_{-\pi}^{\pi} \frac{q(\mathbf{h}(\psi))}{\|\mathbf{g}((\mathbf{F}\,|_{S^1})^{-1}(\mathbf{h}(\psi)))\|}\, d\psi \\ &= \int_{-\pi}^{\pi} \frac{q(\mathbf{h}(\psi))}{\gamma(1-(\mathbf{F}\,|_{S^1})^{-1}(\mathbf{h}(\psi))\cdot\mathbf{v})}\, d\psi.\end{aligned} \tag{32}$$

Since the inverse of the restriction of $\mathbf{F}$ to S^1 is easily seen to be given by the equation

$$(\mathbf{F}\,|_{S^1})^{-1}(\mathbf{r}) = \frac{\mathbf{A}^t\mathbf{r} + (\gamma-1)\mathbf{P}_{-\mathbf{v}}(\mathbf{A}^t\mathbf{r}) + \gamma\mathbf{v}}{\gamma(1+\mathbf{A}^t\mathbf{r}\cdot\mathbf{v})}, \tag{33}$$

we may infer that

$$\begin{aligned}\int_{-\pi}^{\pi} f(\theta)\, d\theta &= \int_{-\pi}^{\pi} \gamma(1+\mathbf{A}^t\mathbf{h}(\psi)\cdot\mathbf{v})q(\mathbf{h}(\psi))\, d\psi \\ &= \gamma\left(1+\int_{-\pi}^{\pi}\mathbf{h}(\psi)\cdot\mathbf{A}\mathbf{v}q(\mathbf{h}(\psi))\, d\psi\right) = \gamma(1+\mathbf{E}_q(\mathbf{r})\cdot\mathbf{A}\mathbf{v}),\end{aligned} \tag{34}$$

where

$$\mathbf{E}_q(\mathbf{r}) := \begin{pmatrix} E_q(x) \\ E_q(y)\end{pmatrix} = \begin{pmatrix} E_q(\cos(\psi)) \\ E_q(\sin(\psi))\end{pmatrix}. \tag{35}$$

Note: the index "q" is inserted here to indicate that the expected values are computed with respect to the distribution described by q (in contradistinction to the distribution described by p).

According to (24) and (34), our transformation equation for probability densities thus assumes the following form:

$$\boxed{p(\mathbf{h}(\theta)) = \frac{q(\mathbf{F}(\mathbf{h}(\theta)))}{\gamma^3(1+\mathbf{E}_q(\mathbf{r})\cdot\mathbf{A}\mathbf{v})(1-\mathbf{h}(\theta)\cdot\mathbf{v})^2}.} \tag{36}$$

In order to convince ourselves that this equation is indeed correct, we will now demonstrate how it can be used to derive the familiar law for the relativistic addition of velocities. Returning again to the situation in Fig. 3 where $\mathbf{w}$ is the velocity of C relative to B, special relativity predicts that the velocity of C, as measured by A, is

$$\mathbf{v} \oplus \mathbf{w} = \frac{\gamma \mathbf{v} + (\gamma - 1)\mathbf{P}_{\mathbf{v}}(\mathbf{A}^t \mathbf{w}) + \mathbf{A}^t \mathbf{w}}{\gamma(1 + \mathbf{A}^t \mathbf{w} \cdot \mathbf{v})}. \tag{37}$$

In light of the general representation of velocities in (15) and (16), we therefore need to show that

$$\mathbf{E}_p(\mathbf{r}) = \frac{\gamma \mathbf{v} + (\gamma - 1)\mathbf{P}_{\mathbf{v}}(\mathbf{A}^t \mathbf{w}) + \mathbf{A}^t \mathbf{w}}{\gamma(1 + \mathbf{A}^t \mathbf{w} \cdot \mathbf{v})}. \tag{38}$$

Since

$$\begin{aligned}\mathbf{E}_p(\mathbf{r}) &= \begin{pmatrix} \int_{-\pi}^{\pi} \cos(\theta) p(\mathbf{h}(\theta))\, d\theta \\ \int_{-\pi}^{\pi} \sin(\theta) p(\mathbf{h}(\theta))\, d\theta \end{pmatrix} = \int_{-\pi}^{\pi} \mathbf{h}(\theta) p(\mathbf{h}(\theta))\, d\theta \\ &= \int_{-\pi}^{\pi} \frac{\mathbf{h}(\theta) q(\mathbf{F}(\mathbf{h}(\theta)))}{\gamma^3 (1 + \mathbf{E}_q(\mathbf{r}) \cdot \mathbf{A}\mathbf{v})(1 - \mathbf{h}(\theta) \cdot \mathbf{v})^2}\, d\theta,\end{aligned} \tag{39}$$

we may use the substitution in (30) in conjunction with (31) and (33) to infer that

$$\begin{aligned}\mathbf{E}_p(\mathbf{r}) &= \int_{-\pi}^{\pi} \frac{(\mathbf{F}\,|_{S^1})^{-1}(\mathbf{h}(\psi)) q(\mathbf{h}(\psi))}{\gamma^2 (1 + \mathbf{E}_q(\mathbf{r}) \cdot \mathbf{A}\mathbf{v})(1 - (\mathbf{F}\,|_{S^1})^{-1}(\mathbf{h}(\psi)) \cdot \mathbf{v})}\, d\psi \\ &= \int_{-\pi}^{\pi} \frac{(\mathbf{A}^t \mathbf{h}(\psi) + (\gamma - 1)\mathbf{P}_{-\mathbf{v}}(\mathbf{A}^t \mathbf{h}(\psi)) + \gamma \mathbf{v}) q(\mathbf{h}(\psi))}{\gamma(1 + \mathbf{E}_q(\mathbf{r}) \cdot \mathbf{A}\mathbf{v})}\, d\psi \\ &= \frac{\mathbf{A}^t \mathbf{E}_q(\mathbf{r}) + (\gamma - 1)\mathbf{P}_{-\mathbf{v}}(\mathbf{A}^t \mathbf{E}_q(\mathbf{r})) + \gamma \mathbf{v}}{\gamma(1 + \mathbf{A}^t \mathbf{E}_q(\mathbf{r}) \cdot \mathbf{v})}.\end{aligned} \tag{40}$$

The desired result in (38) is now an immediate consequence of the fact that $\mathbf{E}_q(\mathbf{r}) = \mathbf{w}$ and that $\mathbf{P}_{-\mathbf{v}}(\mathbf{u}) = \mathbf{P}_{\mathbf{v}}(\mathbf{u})$ for all $\mathbf{u}$. Given this successful application of our formalism, it seems justified to assert that our probabilistic model correctly describes relativistic motion in two spatial dimensions.

Interestingly, in our derivation of (38) we never had to actually find an explicit formula for either p or q—the transformation equation (36) in itself was sufficient. However, an explicit formula is easily obtained if we make the entirely plausible assumption that q describes a uniform distribution in the case where $\mathbf{w}$ is zero. In other words, for $\mathbf{w} = \mathbf{0}$ we set

$$q(\mathbf{h}(\theta)) := \frac{1}{2\pi} \tag{41}$$

for all $\theta \in [-\pi, \pi]$.

Remark. The condition $\mathbf{w} = \mathbf{0}$ does not imply with necessity that the distribution described by q is uniform, for there obviously are infinitely many probability densities on the unit circle that satisfy the equations $E(\cos(\theta)) = E(\sin(\theta)) = 0$ (see (15) and (16)). In fact, it is one of the strengths of our model that it provides this freedom, because flexibility in choosing densities is potentially useful when adjustments to external constraints are needed.

Since the definition of q in (41) clearly implies that $\mathbf{E}_q(\mathbf{r}) = \mathbf{0}$, we may apply (36) with $1/(2\pi)$ in place of q to infer that the probability density associated with a uniform linear motion in the xy-plane is

$$\boxed{p(\mathbf{h}(\theta)) = \frac{1}{2\pi\gamma^3(1 - \mathbf{h}(\theta)\cdot\mathbf{v})^2} = \frac{\sqrt{1 - \|\mathbf{v}\|^2}^{\,3}}{2\pi(1 - \mathbf{h}(\theta)\cdot\mathbf{v})^2}.} \tag{42}$$

That the velocity associated with p is indeed $\mathbf{v} = (v_1, v_2)$ is easily verified: using elementary methods of integration, we find that

$$\mathbf{E}_p(\mathbf{r}) = \begin{pmatrix} \displaystyle\int_{-\pi}^{\pi} \frac{\cos(\theta)}{2\pi\gamma^3(1 - v_1\cos(\theta) - v_2\sin(\theta))^2}\, d\theta \\ \displaystyle\int_{-\pi}^{\pi} \frac{\sin(\theta)}{2\pi\gamma^3(1 - v_1\cos(\theta) - v_2\sin(\theta))^2}\, d\theta \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \tag{43}$$

as desired. Furthermore, using the polar representation $\mathbf{v} = v\mathbf{h}(\alpha)$ for some $\alpha \in [-\pi, \pi]$, the decrease in randomness (that is, in proper time) as v approaches the speed of light $c = 1$ is nicely expressed in the fact that

$$\lim_{v\to 1} p(\mathbf{h}(\theta)) = \lim_{v\to 1} \frac{\sqrt{1 - v^2}^{\,3}}{2\pi(1 - v\mathbf{h}(\theta)\cdot\mathbf{h}(\alpha))^2} = \begin{cases} 0 & \text{for } \theta \neq \alpha, \\ \infty & \text{for } \theta = \alpha. \end{cases} \tag{44}$$

In other words, for $v = 1$, the density p contracts to the Dirac delta function centered at α, and the variance of both $\cos(\theta)$ and $\sin(\theta)$ is zero. A graphical illustration of this limiting behavior for the special case $\alpha = 0$ is shown in Fig. 4.

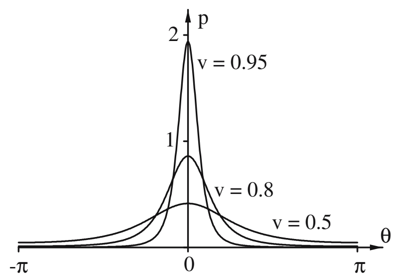

Fig. 4. Graphs of the density p for increasing values of v.

As further evidence for the adequacy of our formalism, we also wish to mention that the form of the density function in (42) is Lorentz invariant. For considering again the situation in Fig. 3, and assuming that the density which B associates with the motion of C is

$$q(\mathbf{h}(\psi)) = \frac{\sqrt{1-\|\mathbf{w}\|^2}^{\,3}}{2\pi(1-\mathbf{h}(\psi)\cdot\mathbf{w})^2}, \tag{45}$$

we may apply (36), (20), (25), and (37) to infer that the corresponding density in the inertial frame of A is

$$\begin{aligned} p(\mathbf{h}(\theta)) &= \frac{\sqrt{1-\|\mathbf{w}\|^2}^{\,3}}{2\pi\gamma^3(1+\mathbf{E}_q(\mathbf{r})\cdot\mathbf{A}\mathbf{v})(1-\mathbf{h}(\theta)\cdot\mathbf{v})^2(1-\mathbf{F}(\mathbf{h}(\theta))\cdot\mathbf{w})^2} \\ &= \frac{\sqrt{1-\|\mathbf{w}\|^2}^{\,3}}{2\pi\gamma^3(1+\mathbf{A}^t\mathbf{w}\cdot\mathbf{v})^3(1-\mathbf{h}(\theta)\cdot\mathbf{v}\oplus\mathbf{w})^2} \\ &= \frac{\sqrt{1-\|\mathbf{v}\oplus\mathbf{w}\|^2}^{\,3}}{2\pi(1-\mathbf{h}(\theta)\cdot\mathbf{v}\oplus\mathbf{w})^2}. \end{aligned} \tag{46}$$

In other words, the density observed by A is of exactly the same form as the one observed by B except that $\mathbf{w}$ is replaced by $\mathbf{v}\oplus\mathbf{w}$ as it should be.

To continue our discussion we will now extend our model to the case of three spatial dimensions. Proceeding in a manner completely analogous to the two-dimensional case, we assume that we are given a probability density p on the unit sphere

$$S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\} \tag{47}$$

such that in spherical coordinates the probability for a move in a random walk to fall into the angular rectangle $[\theta, \theta + d\theta] \times [\phi, \phi + d\phi]$ is

$$p(\mathbf{h}(\theta, \phi)) \sin(\phi)\, d\theta d\phi \tag{48}$$

where

$$\mathbf{h}(\theta,\phi) := \begin{pmatrix} x(\theta,\phi) \\ y(\theta,\phi) \\ z(\theta,\phi) \end{pmatrix} := \begin{pmatrix} \cos(\theta)\sin(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\phi) \end{pmatrix} \tag{49}$$

for all $(\theta, \phi) \in [-\pi, \pi] \times [0, \pi]$. Given this setup, the velocity associated with p is

$$\mathbf{v} = \begin{pmatrix} E(x) \\ E(y) \\ E(z) \end{pmatrix} = \begin{pmatrix} \int_0^\pi \int_{-\pi}^{\pi} x(\theta,\phi)\, p(\mathbf{h}(\theta,\phi)) \sin(\phi)\, d\theta d\phi \\ \int_0^\pi \int_{-\pi}^{\pi} y(\theta,\phi)\, p(\mathbf{h}(\theta,\phi)) \sin(\phi)\, d\theta d\phi \\ \int_0^\pi \int_{-\pi}^{\pi} z(\theta,\phi)\, p(\mathbf{h}(\theta,\phi)) \sin(\phi)\, d\theta d\phi \end{pmatrix}, \tag{50}$$

and the random measure of proper time is

$$\sqrt{1 - \|\mathbf{v}\|^2} = \sqrt{\mathrm{Var}(x) + \mathrm{Var}(y) + \mathrm{Var}(z)}. \tag{51}$$

To address the problem of transforming densities in three dimensions, as we pass from one inertial frame to another, we simply add a z-axis and a ζ-axis to the frames of A and B in Fig. 3, respectively. The corresponding transformation equation for displacement vectors in random walks is the same as in (20), except for the fact that all vectors are now in $\mathbb{R}^3$ and that $\mathbf{A}$ is a 3×3 matrix:

$$\begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \mathbf{A}(\mathbf{r} + (\gamma - 1)\mathbf{P}_{\mathbf{v}}(\mathbf{r}) - \gamma \|\mathbf{r}\| \mathbf{v}) =: \mathbf{g}(\mathbf{r}). \tag{52}$$

Again denoting by q the density that B associates with the motion of C, it is not difficult to see that, in strict analogy to (24), the density p in the inertial frame of A is given by the equation

$$p(\mathbf{h}(\theta,\phi)) \sin(\phi) = \frac{f(\theta,\phi)}{\int_0^\pi \int_{-\pi}^{\pi} f(\theta,\phi)\, d\theta d\phi} \tag{53}$$

where

$$\begin{aligned} f(\theta,\phi) :&= \frac{q\left(\dfrac{\mathbf{g}(\mathbf{h}(\theta,\phi))}{\|\mathbf{g}(\mathbf{h}(\theta,\phi))\|}\right) \left\| \dfrac{\partial}{\partial\theta} \dfrac{\mathbf{g}(\mathbf{h}(\theta,\phi))}{\|\mathbf{g}(\mathbf{h}(\theta,\phi))\|} \times \dfrac{\partial}{\partial\phi} \dfrac{\mathbf{g}(\mathbf{h}(\theta,\phi))}{\|\mathbf{g}(\mathbf{h}(\theta,\phi))\|} \right\|}{\|\mathbf{g}(\mathbf{h}(\theta,\phi))\|} \\ &= \frac{q(\mathbf{F}(\mathbf{h}(\theta,\phi))) \| D_\theta \mathbf{F}(\mathbf{h}(\theta,\phi)) \times D_\phi \mathbf{F}(\mathbf{h}(\theta,\phi)) \|}{\gamma(1 - \mathbf{h}(\theta,\phi) \cdot \mathbf{v})}. \end{aligned} \tag{54}$$

Given the results in (27) and (28), it follows that

$$D_\theta \mathbf{F}(\mathbf{h}(\theta,\phi)) = \frac{\mathbf{A} D_\theta \mathbf{h}(\theta,\phi)}{\gamma(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})} + \frac{D_\theta \mathbf{h}(\theta,\phi)\cdot\mathbf{v}}{(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})^2}\left(\frac{\mathbf{A}\mathbf{h}(\theta,\phi)}{\gamma} - \frac{\mathbf{A}\mathbf{v}}{\gamma+1}\right),$$
$$D_\phi \mathbf{F}(\mathbf{h}(\theta,\phi)) = \frac{\mathbf{A} D_\phi \mathbf{h}(\theta,\phi)}{\gamma(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})} + \frac{D_\phi \mathbf{h}(\theta,\phi)\cdot\mathbf{v}}{(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})^2}\left(\frac{\mathbf{A}\mathbf{h}(\theta,\phi)}{\gamma} - \frac{\mathbf{A}\mathbf{v}}{\gamma+1}\right), \tag{55}$$

and

$$\|D_\theta \mathbf{F}(\mathbf{h}(\theta,\phi))\| = \frac{\|D_\theta \mathbf{h}(\theta,\phi)\|}{\gamma(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})} = \frac{\sin(\phi)}{\gamma(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})},$$
$$\|D_\phi \mathbf{F}(\mathbf{h}(\theta,\phi))\| = \frac{\|D_\phi \mathbf{h}(\theta,\phi)\|}{\gamma(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})} = \frac{1}{\gamma(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})}. \tag{56}$$

Consequently,

$$\begin{aligned}
&\|D_\theta \mathbf{F}(\mathbf{h}(\theta,\phi)) \times D_\phi \mathbf{F}(\mathbf{h}(\theta,\phi))\|^2 \\
&\quad = \|D_\theta \mathbf{F}(\mathbf{h}(\theta,\phi))\|^2 \|D_\phi \mathbf{F}(\mathbf{h}(\theta,\phi))\|^2 - (D_\theta \mathbf{F}(\mathbf{h}(\theta,\phi)) \cdot D_\phi \mathbf{F}(\mathbf{h}(\theta,\phi)))^2 \\
&\quad = \|D_\theta \mathbf{F}(\mathbf{h}(\theta,\phi))\|^2 \|D_\phi \mathbf{F}(\mathbf{h}(\theta,\phi))\|^2 - 0 \\
&\quad = \frac{\sin^2(\phi)}{\gamma^4(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})^4}
\end{aligned} \tag{57}$$

and, by implication,

$$f(\theta,\phi) = \frac{q(\mathbf{F}(\mathbf{h}(\theta,\phi)))\sin(\phi)}{\gamma^3(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})^3}. \tag{58}$$

In order to evaluate the integral of $f(\theta,\phi)$, we introduce the substitution

$$\mathbf{h}(\psi,\varphi) = \mathbf{F}(\mathbf{h}(\theta,\phi)) \tag{59}$$

so that

$$\begin{aligned}
(D_\psi \mathbf{h}(\psi,\varphi), D_\varphi \mathbf{h}(\psi,\varphi)) &= D\mathbf{h}(\psi,\varphi) \\
&= D\mathbf{F}(\mathbf{h}(\theta,\phi)) D\mathbf{h}(\theta,\phi) \begin{pmatrix} \partial\theta/\partial\psi & \partial\theta/\partial\varphi \\ \partial\phi/\partial\psi & \partial\phi/\partial\varphi \end{pmatrix} \\
&= (D_\theta \mathbf{F}(\mathbf{h}(\theta,\phi)), D_\phi \mathbf{F}(\mathbf{h}(\theta,\phi))) \begin{pmatrix} \partial\theta/\partial\psi & \partial\theta/\partial\varphi \\ \partial\phi/\partial\psi & \partial\phi/\partial\varphi \end{pmatrix}.
\end{aligned} \tag{60}$$

Since this equation, in conjunction with (57), is easily seen to imply[3] that

$$\left|\det\begin{pmatrix}\partial\theta/\partial\psi & \partial\theta/\partial\varphi\\ \partial\phi/\partial\psi & \partial\phi/\partial\varphi\end{pmatrix}\right| = \frac{\|D_\psi\mathbf{h}(\psi,\varphi)\times D_\varphi\mathbf{h}(\psi,\varphi)\|}{\|D_\theta\mathbf{F}(\mathbf{h}(\theta,\phi))\times D_\phi\mathbf{F}(\mathbf{h}(\theta,\phi))\|}$$
$$= \frac{\gamma^2(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})^2\sin(\varphi)}{\sin(\phi)}, \tag{61}$$

we may again use (33) (which is valid in two as well as in three spatial dimensions) to infer that

$$\int_0^\pi\int_{-\pi}^\pi f(\theta,\phi)\,d\theta d\phi = \int_0^\pi\int_{-\pi}^\pi \frac{q(\mathbf{h}(\psi,\varphi))\sin(\varphi)}{\gamma(1-(\mathbf{F}\mid_{S^2})^{-1}(\mathbf{h}(\psi,\varphi))\cdot\mathbf{v})}\,d\psi d\varphi$$
$$= \int_0^\pi\int_{-\pi}^\pi \gamma(1+\mathbf{A}^t\mathbf{h}(\psi,\varphi)\cdot\mathbf{v})q(\mathbf{h}(\psi,\varphi))\sin(\varphi)\,d\psi d\varphi$$
$$= \gamma(1+\mathbf{E}_q(\mathbf{r})\cdot\mathbf{A}\mathbf{v}). \tag{62}$$

Combining this result with (53) and (58), we obtain

$$\boxed{p(\mathbf{h}(\theta,\phi)) = \frac{q(\mathbf{F}(\mathbf{h}(\theta,\phi)))}{\gamma^4(1+\mathbf{E}_q(\mathbf{r})\cdot\mathbf{A}\mathbf{v})(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})^3}.} \tag{63}$$

As we apply this formula to the uniform density $q(\mathbf{h}(\psi,\varphi)) = 1/4\pi$, we arrive at the conclusion that a uniform linear motion in three spatial dimensions is described by the density function

$$\boxed{p(\mathbf{h}(\theta,\phi)) = \frac{1}{4\pi\gamma^4(1-\mathbf{h}(\theta,\phi)\cdot\mathbf{v})^3}.} \tag{64}$$

In analogy to Fig. 4, the decrease in randomness as $\|\mathbf{v}\|$ approaches the speed of light is illustrated in Figs. 5, 6, and 7 where three graphs of $p(\mathbf{h}(\theta,\phi))\sin(\phi)$ are shown on $[-\pi,\pi]\times[0,\pi]$ for $\mathbf{v} = (0.5, 0, 0)$, $\mathbf{v} = (0.7, 0, 0)$, and $\mathbf{v} = (0.9, 0, 0)$, respectively. Furthermore, just as in the two-dimensional case, the law stated in (37) concerning the relativistic addition of velocities can be easily derived from the transformation equation (63), and the form of the density function in (64) can be shown to be Lorentz invariant in essentially the same manner as in (46).

[3] If $\mathbf{B}$ is a 2×2 matrix and $\mathbf{v}$, $\mathbf{w}$, $\mathbf{x}$, $\mathbf{y}$ are column vectors in $\mathbb{R}^3$, then the matrix equation $(\mathbf{v},\mathbf{w}) = (\mathbf{x},\mathbf{y})\mathbf{B}$ implies that $\|\mathbf{v}\times\mathbf{w}\|^2 = \|\mathbf{x}\times\mathbf{y}\|^2\det(\mathbf{B})^2$.

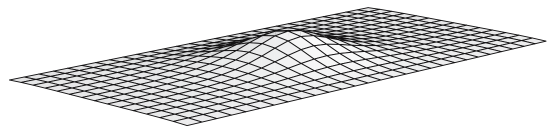

Fig. 5. Graph of $p(\mathbf{h}(\theta, \phi)) \sin(\phi)$ for $\mathbf{v} = (0.5, 0, 0)$.

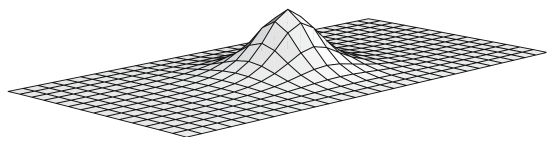

Fig. 6. Graph of $p(\mathbf{h}(\theta, \phi)) \sin(\phi)$ for $\mathbf{v} = (0.7, 0, 0)$.

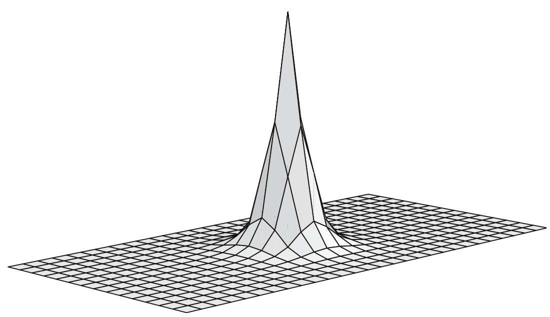

Fig. 7. Graph of $p(\mathbf{h}(\theta, \phi)) \sin(\phi)$ for $\mathbf{v} = (0.9, 0, 0)$.

3. COMPATIBILITY WITH GENERAL RELATIVITY

As we now move on to discuss general relativity, we need to make a few adjustments. To begin with, an observer's proper time will no longer be identical with the process index i that signifies the different stages in a random walk. Instead we will adopt the familiar four-dimensional formalism whereby events in space–time are identified by generalized coordinates $\mathbf{q} = (q^1, q^2, q^3, q^4) \in \mathbb{R}^4$. In order to describe a random walk in such a setting, we further assume that for each triple $(k, j, l) \in \{1, 2, 3, 4\}^3$ we are given a continuously differentiable function $\Gamma^k_{jl} : \mathbb{R}^4 \to \mathbb{R}$. Since this choice of notation is obviously highly suggestive, it is important to point out, that we are not operating within the conceptual framework of differential geometry at this stage—manifolds, tangent spaces, or geodesics are nowhere in sight. As a matter of course, the functions Γ^k_{jl} will eventually assume the role of the familiar Christoffel symbols, but for the moment, our sole concern is to create a framework within which generalized random walks can be described.

With this purpose in mind, let us suppose we are given two sequences of i.i.d. random variables $\theta_0, \ldots, \theta_{n-1} \in [-\pi, \pi]$ and $\phi_0, \ldots, \phi_{n-1} \in [0, \pi]$ such that the joint density of each pair (θ_i, ϕ_i) is

$$p(\mathbf{h}(\theta, \phi)) \sin(\phi) = \frac{\sin(\phi)}{4\pi\gamma^4(1 - \mathbf{h}(\theta, \phi) \cdot \mathbf{v})^3} \tag{65}$$

for some fixed vector $\mathbf{v} = (v^1, v^2, v^3) \in \mathbb{R}^3$ with $\|\mathbf{v}\| \leqslant 1$. Setting

$$\begin{aligned} x_i &:= \cos(\theta_i) \sin(\phi_i), \\ y_i &:= \sin(\theta_i) \sin(\phi_i), \\ z_i &:= \cos(\phi_i), \end{aligned} \tag{66}$$

we observe that each of the sequences $x_0, \ldots, x_{n-1}$, $y_0, \ldots, y_{n-1}$, and $z_0, \ldots, z_{n-1}$ consists of i.i.d. random variables as well. (Note: independence is here asserted only for random variables with pairwise distinct indices.) Moreover, according to (50), we may infer that

$$\mathbf{v} = \begin{pmatrix} E(x_i) \\ E(y_i) \\ E(z_i) \end{pmatrix} =: \begin{pmatrix} \mu_x \\ \mu_y \\ \mu_z \end{pmatrix}. \tag{67}$$

In order to construct a random walk in the q^k-coordinate system, we choose a step-length parameter $\Delta t > 0$, an initial point $\mathbf{q}_0$, and a set of vectors $\{\mathbf{T}_0, \mathbf{X}_0, \mathbf{Y}_0, \mathbf{Z}_0\}$. Since random walks in the current context are most appropriately described recursively, we also assume that we are given a point $\mathbf{q}_i$ and vectors $\mathbf{T}_i = (T_i^k)$, $\mathbf{X}_i = (X_i^k)$, $\mathbf{Y}_i = (Y_i^k)$, and $\mathbf{Z}_i = (Z_i^k)$ for some $i \in \{0, \ldots, n-1\}$. Then, using Einstein's summation convention (with respect to the indices j and l), we set

$$\mathbf{q}_{i+1} := \mathbf{q}_i + (\mathbf{T}_i + x_i \mathbf{X}_i + y_i \mathbf{Y}_i + z_i \mathbf{Z}_i)\Delta t, \tag{68}$$

$$\begin{aligned} T_{i+1}^k &:= T_i^k - \Gamma_{jl}^k(\mathbf{q}_i)(T_i^l + x_i X_i^l + y_i Y_i^l + z_i Z_i^l) T_i^j \Delta t, \\ X_{i+1}^k &:= X_i^k - \Gamma_{jl}^k(\mathbf{q}_i)(T_i^l + x_i X_i^l + y_i Y_i^l + z_i Z_i^l) X_i^j \Delta t, \\ Y_{i+1}^k &:= Y_i^k - \Gamma_{jl}^k(\mathbf{q}_i)(T_i^l + x_i X_i^l + y_i Y_i^l + z_i Z_i^l) Y_i^j \Delta t, \\ Z_{i+1}^k &:= Z_i^k - \Gamma_{jl}^k(\mathbf{q}_i)(T_i^l + x_i X_i^l + y_i Y_i^l + z_i Z_i^l) Z_i^j \Delta t. \end{aligned} \tag{69}$$

*Remark***.** The random walks of special relativity, as discussed in Sec. 2, are obtained from these defining equations if we set

$$\begin{aligned} &\Delta t = 1, \quad \Gamma_{jl}^k(\mathbf{q}) = 0, \quad \mathbf{T}_0 = (1, 0, 0, 0), \\ &\mathbf{X}_0 = (0, 1, 0, 0), \;\; \mathbf{Y}_0 = (0, 0, 1, 0), \quad \mathbf{Z}_0 = (0, 0, 0, 1). \end{aligned} \tag{70}$$

Furthermore, in light of the anticipated interpretation of the functions Γ^k_{jl} as the Christoffel symbols of a space–time manifold, the equations listed in (69) are easily seen to describe the discrete "parallel shifts" (intuitively speaking) of the vectors $\mathbf{T}_0$, $\mathbf{X}_0$, $\mathbf{Y}_0$, and $\mathbf{Z}_0$ along the random walk given by the points $\mathbf{q}_0, \ldots, \mathbf{q}_n$. To understand why this is so, let us consider the first step from $\mathbf{q}_0$ to $\mathbf{q}_1$ and the corresponding shift of $\mathbf{T}_0$ to $\mathbf{T}_1$. The step itself consists of the "spatial" and "temporal" displacements $(x_0\mathbf{X}_0 + y_0\mathbf{Y}_0 + z_0\mathbf{Z}_0)\Delta t$ and $\mathbf{T}_0\Delta t$, respectively. Writing the parallel transport equation

$$\frac{dV^k}{dt} + \Gamma^k_{jl} V^j \frac{dq^l}{dt} = 0 \tag{71}$$

as a difference equation yields

$$V^k(t + \Delta t) = V^k(t) - \Gamma^k_{jl} V^j(t)\Delta q^l. \tag{72}$$

Replacing in this order $V^k(t+\Delta t)$, $V^k(t)$, and Δq^l with T^k_1, T^k_0, and $(T^l_0 + x_0 X^l_0 + y_0 Y^l_0 + z_0 Z^l_0)\Delta t$, we notice that the approximate parallel shift of the vector $\mathbf{T}_0$ along the step from $\mathbf{q}_0$ to $\mathbf{q}_1$ is given by the equation

$$T^k_1 = T^k_0 - \Gamma^k_{jl}(\mathbf{q}_0)(T^l_0 + x_0 X^l_0 + y_0 Y^l_0 + z_0 Z^l_0)T^j_0\,\Delta t. \tag{73}$$

In other words, it is given by the first equation in (69) for $i = 0$.

Our goal for the remainder of this section will be to demonstrate that

(i) any random walk consisting of a sequence of points $\mathbf{q}_0, \ldots, \mathbf{q}_n$, as described in (68) and (69), contracts with probability 1 to a curve in the q^k-coordinate system as Δt approaches zero (and as n tends to infinity), and that

(ii) this curve is a geodesic if the functions Γ^k_{jl} are defined to be the Christoffel symbols of a space–time manifold.

It is our view that in establishing the validity of these claims, our probabilistic model of motion will have been shown to be compatible—at least in a very basic sense—with the standard continuous model in the statistical limit.

Before we proceed with the proof of (i), it is appropriate that we simplify our notation: for all $i \in \{0, \ldots, n-1\}$ we define a 4×4 matrix $\mathbf{B}_i$ by the equation

$$\mathbf{B}_i := -(\Gamma^k_{jl}(\mathbf{q}_i)(T^l_i + x_i X^l_i + y_i Y^l_i + z_i Z^l_i))^4_{k,j=1} \tag{74}$$

where k is the row index and j the column index. Given this definition, it is easy to verify that

$$\begin{aligned}
\mathbf{T}_i &= \left(\prod_{m=0}^{i-1}(\mathbf{I}+\mathbf{B}_m\Delta t)\right)\mathbf{T}_0,\\
\mathbf{X}_i &= \left(\prod_{m=0}^{i-1}(\mathbf{I}+\mathbf{B}_m\Delta t)\right)\mathbf{X}_0,\\
\mathbf{Y}_i &= \left(\prod_{m=0}^{i-1}(\mathbf{I}+\mathbf{B}_m\Delta t)\right)\mathbf{Y}_0,\\
\mathbf{Z}_i &= \left(\prod_{m=0}^{i-1}(\mathbf{I}+\mathbf{B}_m\Delta t)\right)\mathbf{Z}_0
\end{aligned}\tag{75}$$

for all $i \in \{1, \ldots, n-1\}$. Since clearly

$$\mathbf{q}_n = \mathbf{q}_0 + \sum_{i=0}^{n-1}(\mathbf{T}_i + x_i\mathbf{X}_i + y_i\mathbf{Y}_i + z_i\mathbf{Z}_i)\Delta t, \tag{76}$$

it will be helpful to take a closer look at the vector

$$\sum_{i=0}^{n-1}\mathbf{T}_i = \mathbf{T}_0 + \sum_{i=1}^{n-1}\left(\prod_{m=0}^{i-1}(\mathbf{I}+\mathbf{B}_m\Delta t)\right)\mathbf{T}_0. \tag{77}$$

Setting

$$\mathbf{D}_i := \prod_{m=0}^{i-1}(\mathbf{I}+\mathbf{B}_m\Delta t) - \mathbf{I}, \tag{78}$$

we obtain

$$\sum_{i=0}^{n-1}\mathbf{T}_i = n\mathbf{T}_0 + \sum_{i=1}^{n-1}\mathbf{D}_i\mathbf{T}_0. \tag{79}$$

In order to find an estimate for the operator norms[4] of the matrices $\mathbf{D}_i$, we will now discuss some technical details that are of little interest in themselves but are necessary if we wish to maintain an adequate level of rigor.

[4] As usual, the operator norm of an $n \times n$ matrix $\mathbf{A}$ is defined as $\|\mathbf{A}\| = \sup\{\|\mathbf{A}\mathbf{x}\| \mid \|\mathbf{x}\| = 1\}$ where $\|\mathbf{x}\|$ denotes the Euclidean norm $(\sum_{k=1}^{n} x_k^2)^{1/2}$.

To begin with, we will require that all random walks are contained in some large compact set $M \subset \mathbb{R}^4$ such as a closed ball with a giant radius. Then the assumed continuous differentiability of the functions Γ^k_{jl} implies in particular that these functions are universally bounded on M (mere continuity would be sufficient for this conclusion, but continuous differentiability will be needed later). Since, as we saw, the vectors $\mathbf{T}_i$, $\mathbf{X}_i$, $\mathbf{Y}_i$, and $\mathbf{Z}_i$ are to be thought of as discrete approximations of parallel shifts along random walks, and since the functions Γ^k_{jl} are bounded, it is reasonable to assume that for all random walks considered from here on the Euclidean norms of the vectors $\mathbf{T}_i$, $\mathbf{X}_i$, $\mathbf{Y}_i$, and $\mathbf{Z}_i$ are universally bounded as well. More precisely, we assume that there is a constant $\nu > 0$ such that

$$\|\mathbf{T}_i\| + \|\mathbf{X}_i\| + \|\mathbf{Y}_i\| + \|\mathbf{Z}_i\| \leqslant \nu \tag{80}$$

for all indices i in all random walks. Given the definition of the matrices $\mathbf{B}_i$ in (74) in terms of the functions Γ^k_{jl} and the components T^k_i, X^k_i, Y^k_i, Z^k_i, we may thus infer that there also exists a constant $\lambda > 0$ such that universally

$$\|\mathbf{B}_i\| \leqslant \lambda. \tag{81}$$

A more rigorous approach whereby the existence of ν and λ is proven to be a consequence of the boundedness of the functions Γ^k_{jl} and of the assumed existence of a bound for the initial lengths $\|\mathbf{T}_0\|$, $\|\mathbf{X}_0\|$, $\|\mathbf{Y}_0\|$, and $\|\mathbf{Z}_0\|$ is feasible but very tedious.

Using (81), it is now easy to see that

$$\|\mathbf{D}_i\| \leqslant \sum_{m=1}^{i} \binom{i}{m} \lambda^m \Delta t^m = (1 + \lambda\Delta t)^i - 1, \tag{82}$$

and therefore,

$$\begin{aligned} \left\| \sum_{i=1}^{n-1} \mathbf{D}_i \mathbf{T}_0 \right\| &\leqslant \sum_{i=1}^{n-1} ((1 + \lambda\Delta t)^i - 1)\|\mathbf{T}_0\| \\ &= \frac{((1 + \lambda\Delta t)^n - 1 - n\lambda\Delta t)\|\mathbf{T}_0\|}{\lambda\Delta t}. \end{aligned} \tag{83}$$

Similarly, we find that

$$\sum_{i=0}^{n-1} x_i \mathbf{X}_i = \sum_{i=0}^{n-1} x_i \mathbf{X}_0 + \sum_{i=1}^{n-1} x_i \mathbf{D}_i \mathbf{X}_0 \tag{84}$$

and

$$\left\| \sum_{i=1}^{n-1} x_i \mathbf{D}_i \mathbf{X}_0 \right\| \leqslant \frac{((1+\lambda\Delta t)^n - 1 - n\lambda\Delta t)\|\mathbf{X}_0\|}{\lambda\Delta t} \quad \text{(because } |x_i| \leqslant 1\text{)}. \tag{85}$$

Since completely analogous estimates are valid for the **Y**- and **Z**-components as well, we may apply (76) to infer that the distance of $\mathbf{q}_n$ from the vector

$$\begin{aligned} \mathbf{r}_n &:= \mathbf{q}_0 + n(\mathbf{T}_0 + v^1\mathbf{X}_0 + v^2\mathbf{Y}_0 + v^3\mathbf{Z}_0)\Delta t \\ &= \mathbf{q}_0 + n(\mathbf{T}_0 + \mu_x\mathbf{X}_0 + \mu_y\mathbf{Y}_0 + \mu_z\mathbf{Z}_0)\Delta t \end{aligned} \tag{86}$$

satisfies the following inequality:

$$\begin{aligned} \|\mathbf{q}_n - \mathbf{r}_n\| &\leqslant \left\| \mathbf{q}_n - \mathbf{q}_0 - \sum_{i=0}^{n-1} (\mathbf{T}_0 + x_i\mathbf{X}_0 + y_i\mathbf{Y}_0 + z_i\mathbf{Z}_0)\Delta t \right\| \\ &\quad + \left\| \mathbf{q}_0 + \sum_{i=0}^{n-1} (\mathbf{T}_0 + x_i\mathbf{X}_0 + y_i\mathbf{Y}_0 + z_i\mathbf{Z}_0)\Delta t - \mathbf{r}_n \right\| \\ &= \left\| \sum_{i=1}^{n-1} (\mathbf{D}_i\mathbf{T}_0 + x_i\mathbf{D}_i\mathbf{X}_0 + y_i\mathbf{D}_i\mathbf{Y}_0 + z_i\mathbf{D}_i\mathbf{Z}_0)\Delta t \right\| \\ &\quad + \left\| \mathbf{q}_0 + \sum_{i=0}^{n-1} (\mathbf{T}_0 + x_i\mathbf{X}_0 + y_i\mathbf{Y}_0 + z_i\mathbf{Z}_0)\Delta t - \mathbf{r}_n \right\| \\ &\leqslant \frac{((1+\lambda\Delta t)^n - 1 - n\lambda\Delta t)(\|\mathbf{T}_0\| + \|\mathbf{X}_0\| + \|\mathbf{Y}_0\| + \|\mathbf{Z}_0\|)}{\lambda} \\ &\quad + n(|\bar{x} - \mu_x|\|\mathbf{X}_0\| + |\bar{y} - \mu_y|\|\mathbf{Y}_0\| + |\bar{z} - \mu_z|\|\mathbf{Z}_0\|)\Delta t \\ &\leqslant \frac{\nu((1+\lambda\Delta t)^n - 1 - n\lambda\Delta t)}{\lambda} \\ &\quad + \nu n(|\bar{x} - \mu_x| + |\bar{y} - \mu_y| + |\bar{z} - \mu_z|)\Delta t \end{aligned} \tag{87}$$

where $\bar{x}$, $\bar{y}$, and $\bar{z}$ are the respective sample means of the random variables x_i, y_i, and z_i (i.e., $\bar{x} = \sum_{i=0}^{n-1} x_i/n$, etc.). Setting

$$\sigma_x := \sqrt{\operatorname{Var}(x_i)}, \quad \sigma_y := \sqrt{\operatorname{Var}(y_i)}, \quad \text{and } \sigma_z := \sqrt{\operatorname{Var}(z_i)}, \tag{88}$$

the pairwise independence and pairwise identical distribution of the random variables x_i implies that $\operatorname{Var}(\bar{x}) = \sigma_x^2/n$. Consequently, Chebyshev's inequality allows us to conclude that the probability for $\bar{x}$ to differ from

$E(\bar{x}) = E(x_i) = \mu_x$ by no more than ε (for any $\varepsilon > 0$) is greater than or equal to $1 - \sigma_x^2/(\varepsilon^2 n)$, i.e.,

$$P(|\bar{x} - \mu_x| \leqslant \varepsilon) \geqslant 1 - \frac{\sigma_x^2}{\varepsilon^2 n}. \tag{89}$$

Similarly, Chebyshev's inequality also implies that

$$\begin{aligned} P(|\bar{y} - \mu_y| \leqslant \varepsilon) &\geqslant 1 - \frac{\sigma_y^2}{\varepsilon^2 n}, \\ P(|\bar{z} - \mu_z| \leqslant \varepsilon) &\geqslant 1 - \frac{\sigma_z^2}{\varepsilon^2 n}. \end{aligned} \tag{90}$$

Combining these estimates with (87) yields

$$P(\|\mathbf{q}_n - \mathbf{r}_n\| \leqslant \rho(n, \Delta t, \varepsilon)) \geqslant 1 - \frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}{\varepsilon^2 n} \tag{91}$$

where

$$\rho(n, \Delta t, \varepsilon) := \frac{\nu((1 + \lambda \Delta t)^n - 1 - n\lambda \Delta t)}{\lambda} + 3\nu n \varepsilon \Delta t. \tag{92}$$

In order to examine how the estimate in (91) is affected as Δt decreases to zero, we choose a fixed value $t > 0$ and set

$$\Delta t := \frac{t}{n}. \tag{93}$$

This choice of Δt guarantees that the position vector $\mathbf{r}_n$ remains constant as n increases to infinity (or, equivalently, as Δt decreases to zero), because the definition in (86) implies that

$$\mathbf{r}_n = \mathbf{q}_0 + t(\mathbf{T}_0 + v^1\mathbf{X}_0 + v^2\mathbf{Y}_0 + v^3\mathbf{Z}_0) =: \mathbf{r}. \tag{94}$$

Furthermore, with regard to the estimate in (91) we observe that

$$\begin{aligned} \lim_{n\to\infty} \rho(n, \Delta t, \varepsilon) &= \lim_{n\to\infty} \left(\frac{\nu((1 + \lambda t/n)^n - 1 - \lambda t)}{\lambda} + 3\nu\varepsilon t \right) \\ &= \frac{\nu(e^{\lambda t} - 1 - \lambda t)}{\lambda} + 3\nu\varepsilon t. \end{aligned} \tag{95}$$

Using Taylor's theorem, we obtain

$$\lim_{n\to\infty} \rho(n, \Delta t, \varepsilon) \leqslant \nu\lambda t^2 e^{\lambda t} + 3\nu\varepsilon t, \tag{96}$$

and therefore,

$$
\begin{aligned}
1 &\geqslant \lim_{n\to\infty} P\left(\|\mathbf{q}_n - \mathbf{r}\| \leqslant \nu\lambda t^2 e^{\lambda t} + 3\nu\varepsilon t\right) \\
&\geqslant \lim_{n\to\infty} P(\|\mathbf{q}_n - \mathbf{r}\| \leqslant \rho(n, \Delta t, \varepsilon)) \\
&\geqslant \lim_{n\to\infty} \left(1 - \frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}{\varepsilon^2 n}\right) = 1.
\end{aligned} \tag{97}
$$

Consequently, we arrive at the following conclusion:

$$
\lim_{n\to\infty} P\left(\|\mathbf{q}_n - \mathbf{r}\| \leqslant \nu\lambda t^2 e^{\lambda t} + 3\nu\varepsilon t\right) = 1. \tag{98}
$$

Having thus found an upper estimate for the statistical uncertainty in the position of $\mathbf{q}_n$, it is crucially important that we now derive similar estimates for the vectors $\mathbf{T}_n$, $\mathbf{X}_n$, $\mathbf{Y}_n$, and $\mathbf{Z}_n$. Considering $\mathbf{T}_n$ first, we define a matrix $\mathbf{V}_0$ for the approximate "parallel transport" of $\mathbf{T}_0$ in the direction of the vector $\mathbf{T}_0 + v^1\mathbf{X}_0 + v^2\mathbf{Y}_0 + v^3\mathbf{Z}_0 = \mathbf{T}_0 + \mu_x\mathbf{X}_0 + \mu_y\mathbf{Y}_0 + \mu_z\mathbf{Z}_0$ via the equation

$$
\mathbf{V}_0 := -(\Gamma^k_{jl}(\mathbf{q}_0)(T_0^l + \mu_x X_0^l + \mu_y Y_0^l + \mu_z Z_0^l))^4_{k,j=1}. \tag{99}
$$

Then it follows that

$$
\begin{aligned}
\|\mathbf{T}_n - (\mathbf{T}_0 + n\mathbf{V}_0\mathbf{T}_0\Delta t)\| &= \left\|\sum_{i=0}^{n-1}(\mathbf{T}_{i+1} - \mathbf{T}_i) - n\mathbf{V}_0\mathbf{T}_0\Delta t\right\| \\
&= \left\|\sum_{i=0}^{n-1}\mathbf{B}_i\mathbf{T}_i - n\mathbf{V}_0\mathbf{T}_0\right\|\Delta t \\
&\leqslant \left\|\sum_{i=0}^{n-1}\mathbf{B}_i(\mathbf{T}_i - \mathbf{T}_0)\right\|\Delta t + \left\|\sum_{i=0}^{n-1}(\mathbf{B}_i - \mathbf{V}_0)\mathbf{T}_0\right\|\Delta t \\
&\leqslant \lambda\sum_{i=0}^{n-1}\|\mathbf{T}_i - \mathbf{T}_0\|\Delta t + \nu\left\|\sum_{i=0}^{n-1}(\mathbf{B}_i - \mathbf{V}_0)\right\|\Delta t \\
&= \lambda\sum_{i=0}^{n-1}\left\|\sum_{m=0}^{i-1}(\mathbf{T}_{m+1} - \mathbf{T}_m)\right\|\Delta t + \nu\left\|\sum_{i=0}^{n-1}(\mathbf{B}_i - \mathbf{V}_0)\right\|\Delta t \\
&\leqslant \lambda\sum_{i=0}^{n-1}\sum_{m=0}^{i-1}\|\mathbf{B}_m\mathbf{T}_m\|\Delta t^2 + \nu\left\|\sum_{i=0}^{n-1}(\mathbf{B}_i - \mathbf{V}_0)\right\|\Delta t \\
&\leqslant n^2\nu\lambda^2\Delta t^2 + \nu\left\|\sum_{i=0}^{n-1}(\mathbf{B}_i - \mathbf{V}_0)\right\|\Delta t.
\end{aligned} \tag{100}
$$

In order to find an estimate for the remaining norm on the right, we define

$$\mathbf{C}_m(i) := -(\Gamma^k_{jl}(\mathbf{q}_m)(T^l_m + x_i X^l_m + y_i Y^l_m + z_i Z^l_m))^4_{k,j=1} \tag{101}$$

for all $m \in \{0, \ldots, i\}$ and all $i \in \{0, \ldots, n-1\}$. In other words, $\mathbf{C}_m(i)$ is equal to $\mathbf{B}_m$ except that the coefficients x_m, y_m, and z_m in $\mathbf{B}_m$ are replaced with x_i, y_i, and z_i, respectively. Given this definition, we see that

$$\begin{aligned}\left\|\sum_{i=0}^{n-1}(\mathbf{B}_i - \mathbf{V}_0)\right\| &= \left\|\sum_{i=1}^{n-1}\sum_{m=0}^{i-1}(\mathbf{C}_{m+1}(i) - \mathbf{C}_m(i)) + \sum_{i=0}^{n-1}(\mathbf{C}_0(i) - \mathbf{V}_0)\right\| \\ &\leqslant \sum_{i=1}^{n-1}\sum_{m=0}^{i-1}\|\mathbf{C}_{m+1}(i) - \mathbf{C}_m(i)\| + \left\|\sum_{i=0}^{n-1}(\mathbf{C}_0(i) - \mathbf{V}_0)\right\|. \end{aligned} \tag{102}$$

Denoting by $b^l_{m,j}$ the elements of $\mathbf{B}_m$, we obtain

$$\begin{aligned}|T^l_{m+1} - T^l_m| &= |b^l_{m,j} T^j_m \Delta t| \leqslant \nu\lambda\Delta t, \\ |x_i(X^l_{m+1} - X^l_m)| &= |x_i b^l_{m,j} X^j_m \Delta t| \leqslant \nu\lambda\Delta t, \\ |y_i(Y^l_{m+1} - Y^l_m)| &= |y_i b^l_{m,j} Y^j_m \Delta t| \leqslant \nu\lambda\Delta t, \\ |z_i(Z^l_{m+1} - Z^l_m)| &= |z_i b^l_{m,j} Z^j_m \Delta t| \leqslant \nu\lambda\Delta t, \end{aligned} \tag{103}$$

and therefore, the definition in (101), implies that

$$\begin{aligned}&\|\mathbf{C}_{m+1}(i) - \mathbf{C}_m(i)\| \\ &\quad\leqslant \|(\Gamma^k_{jl}(\mathbf{q}_{m+1})(T^l_{m+1} - T^l_m))^4_{k,j=1}\| \\ &\qquad + \|(\Gamma^k_{jl}(\mathbf{q}_{m+1})x_i(X^l_{m+1} - X^l_m))^4_{k,j=1}\| \\ &\qquad + \|(\Gamma^k_{jl}(\mathbf{q}_{m+1})y_i(Y^l_{m+1} - Y^l_m))^4_{k,j=1}\| \\ &\qquad + \|(\Gamma^k_{jl}(\mathbf{q}_{m+1})z_i(Z^l_{m+1} - Z^l_m))^4_{k,j=1}\| \\ &\qquad + \|((\Gamma^k_{jl}(\mathbf{q}_{m+1}) - \Gamma^k_{jl}(\mathbf{q}_m))(T^l_m + x_i X^l_m + y_i Y^l_m + z_i Z^l_m))^4_{k,j=1}\| \\ &\quad\leqslant \sum_{l=1}^{4}\|(\Gamma^k_{jl}(\mathbf{q}_{m+1}))^4_{k,j=1}\||T^l_{m+1} - T^l_m| \\ &\qquad + \sum_{l=1}^{4}\|(\Gamma^k_{jl}(\mathbf{q}_{m+1}))^4_{k,j=1}\||x_i(X^l_{m+1} - X^l_m)| \\ &\qquad + \sum_{l=1}^{4}\|(\Gamma^k_{jl}(\mathbf{q}_{m+1}))^4_{k,j=1}\||y_i(Y^l_{m+1} - Y^l_m)| \end{aligned}$$

$$+\sum_{l=1}^{4}\|(\Gamma^k_{jl}(\mathbf{q}_{m+1}))^4_{k,j=1}\||z_i(Z^l_{m+1}-Z^l_m)|$$
$$+\sum_{l=1}^{4}\|(\Gamma^k_{jl}(\mathbf{q}_{m+1})-\Gamma^k_{jl}(\mathbf{q}_m))^4_{k,j=1}\||T^l_m+x_iX^l_m+y_iY^l_m+z_iZ^l_m|$$
$$\leqslant 4\nu\lambda\Delta t\sum_{l=1}^{4}\|(\Gamma^k_{jl}(\mathbf{q}_{m+1}))^4_{k,j=1}\|+\nu\sum_{l=1}^{4}\|(\Gamma^k_{jl}(\mathbf{q}_{m+1})-\Gamma^k_{jl}(\mathbf{q}_m))^4_{k,j=1}\|. \tag{104}$$

Since the functions Γ^k_{jl} are continuously differentiable, and since M is compact, the derivative of each Γ^k_{jl} must be bounded on M. Therefore, the mean value theorem for multivariable functions allows us to infer the existence of a constant $D > 0$ such that

$$\|(\Gamma^k_{jl}(\mathbf{p})-\Gamma^k_{jl}(\mathbf{q}))^4_{k,j=1}\|\leqslant D\|\mathbf{p}-\mathbf{q}\| \tag{105}$$

for all $\mathbf{p},\mathbf{q} \in M$. Furthermore, since the functions Γ^k_{jl} themselves are bounded on M as well, we may assume the constant D to be so large that also

$$\|(\Gamma^k_{jl}(\mathbf{q}))^4_{k,j=1}\|\leqslant D \tag{106}$$

for all $\mathbf{q}\in M$ and all indices j, l, and k. Hence

$$\begin{aligned}\|\mathbf{C}_{m+1}(i)-\mathbf{C}_m(i)\| &\leqslant 16D\nu\lambda\Delta t+4D\nu\|\mathbf{q}_{m+1}-\mathbf{q}_m\|\\ &= 16D\nu\lambda\Delta t+4D\nu\|\mathbf{T}_m+x_m\mathbf{X}_m+y_m\mathbf{Y}_m+z_m\mathbf{Z}_m\|\Delta t\\ &\leqslant 16D\nu(\lambda+\nu)\Delta t. \end{aligned}\tag{107}$$

Looking back at (102), we further observe that

$$\begin{aligned}&\left\|\sum_{i=0}^{n-1}(\mathbf{C}_0(i)-\mathbf{V}_0)\right\|\\ &\quad=\left\|\sum_{i=0}^{n-1}(\Gamma^k_{jl}(\mathbf{q}_0)((x_i-\mu_x)X^l_0+(y_i-\mu_y)Y^l_0+(z_i-\mu_z)Z^l_0))^4_{k,j=1}\right\|\\ &\quad=\left\|(\Gamma^k_{jl}(\mathbf{q}_0)(n(\bar{x}-\mu_x)X^l_0+n(\bar{y}-\mu_y)Y^l_0+n(\bar{z}-\mu_z)Z^l_0))^4_{k,j=1}\right\|\\ &\quad\leqslant n\sum_{l=1}^{4}\left\|(\Gamma^k_{jl}(\mathbf{q}_0))^4_{k,j=1}\right\||(\bar{x}-\mu_x)X^l_0+(\bar{y}-\mu_y)Y^l_0+(\bar{z}-\mu_z)Z^l_0|\\ &\quad\leqslant 4n\nu D(|\bar{x}-\mu_x|+|\bar{y}-\mu_y|+|\bar{z}-\mu_z|).\end{aligned}\tag{108}$$

Combining this result with (100), (102), and (107), we finally conclude that

$$\|\mathbf{T}_n - (\mathbf{T}_0 + n\mathbf{V}_0\mathbf{T}_0\Delta t)\| \leqslant n^2\nu\lambda^2\Delta t^2 + 16n^2\nu^2 D(\lambda+\nu)\Delta t^2 + 4n\nu^2 D(|\bar{x}-\mu_x| + |\bar{y}-\mu_y| + |\bar{z}-\mu_z|)\Delta t. \tag{109}$$

Completely analogous estimates are valid for $\mathbf{X}_n$, $\mathbf{Y}_n$, and $\mathbf{Z}_n$ as well:

$$\begin{aligned}
\|\mathbf{X}_n - (\mathbf{X}_0 + n\mathbf{V}_0\mathbf{X}_0\Delta t)\| &\leqslant n^2\nu\lambda^2\Delta t^2 + 16n^2\nu^2 D(\lambda+\nu)\Delta t^2 \\
&\quad + 4n\nu^2 D(|\bar{x}-\mu_x| + |\bar{y}-\mu_y| + |\bar{z}-\mu_z|)\Delta t, \\
\|\mathbf{Y}_n - (\mathbf{Y}_0 + n\mathbf{V}_0\mathbf{Y}_0\Delta t)\| &\leqslant n^2\nu\lambda^2\Delta t^2 + 16n^2\nu^2 D(\lambda+\nu)\Delta t^2 \\
&\quad + 4n\nu^2 D(|\bar{x}-\mu_x| + |\bar{y}-\mu_y| + |\bar{z}-\mu_z|)\Delta t, \\
\|\mathbf{Z}_n - (\mathbf{Z}_0 + n\mathbf{V}_0\mathbf{Z}_0\Delta t)\| &\leqslant n^2\nu\lambda^2\Delta t^2 + 16n^2\nu^2 D(\lambda+\nu)\Delta t^2 \\
&\quad + 4n\nu^2 D(|\bar{x}-\mu_x| + |\bar{y}-\mu_y| + |\bar{z}-\mu_z|)\Delta t.
\end{aligned} \tag{110}$$

As we now turn our attention to the problem of establishing the validity of claim (i), we will assume that in speaking of the contraction of a random walk to a curve in the q^k-coordinate system we are by definition asserting the statistical uncertainty in the position of $\mathbf{q}_n$ for $n = t/\Delta t$ to converge to zero. More precisely, we are asserting that

$$\lim_{n\to\infty} P\left(\|\mathbf{q}_n - \mathbf{r}(t)\| \leqslant \delta\right) = 1 \text{ for all } \delta > 0 \text{ and some } \mathbf{r}(t) \in \mathbb{R}^4. \tag{111}$$

In order to prove this statement, we will break up a given random walk of length n into a finite number of segments and iteratively apply the result in (98) to each of them. To begin with, we pick an integer $k > 0$ and a value $\varepsilon > 0$ such that for a given $\delta > 0$ the following conditions are satisfied:

$$\begin{aligned}
&\frac{\nu\lambda t^2 e^{\lambda t/k}}{k} + 3\nu\varepsilon t < \delta, \\
&\frac{4t^2\nu\lambda^2 + 64t^2\nu^2(\lambda+\nu)}{k} < \delta.
\end{aligned} \tag{112}$$

Having thus chosen k and ε, we will now define a discrete approximation for the curve that $\mathbf{r}(t)$ supposedly is located on. Proceeding in a manner analogous to (68) and (69), we set $\widetilde{\mathbf{q}}_0 := \mathbf{q}_0$, $\widetilde{\mathbf{T}}_0 := \mathbf{T}_0$, $\widetilde{\mathbf{X}}_0 := \mathbf{X}_0$, $\widetilde{\mathbf{Y}}_0 := \mathbf{Y}_0$, $\widetilde{\mathbf{Z}}_0 := \mathbf{Z}_0$, and

$$\begin{aligned}
\widetilde{\mathbf{q}}_{i+1} &:= \widetilde{\mathbf{q}}_i + \frac{t}{k}(\widetilde{\mathbf{T}}_i + v^1\widetilde{\mathbf{X}}_i + v^2\widetilde{\mathbf{Y}}_i + v^3\widetilde{\mathbf{Z}}_i), \\
\widetilde{T}^h_{i+1} &:= \widetilde{T}^h_i - \frac{t}{k}\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)(\widetilde{T}^l_i + v^1\widetilde{X}^l_i + v^2\widetilde{Y}^l_i + v^3\widetilde{Z}^l_i)\widetilde{T}^j_i,
\end{aligned}$$

$$\widetilde{X}^h_{i+1} := \widetilde{X}^h_i - \frac{t}{k}\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)(\widetilde{T}^l_i + v^1\widetilde{X}^l_i + v^2\widetilde{Y}^l_i + v^3\widetilde{Z}^l_i)\widetilde{X}^j_i,$$
$$\widetilde{Y}^h_{i+1} := \widetilde{Y}^h_i - \frac{t}{k}\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)(\widetilde{T}^l_i + v^1\widetilde{X}^l_i + v^2\widetilde{Y}^l_i + v^3\widetilde{Z}^l_i)\widetilde{Y}^j_i,$$
$$\widetilde{Z}^h_{i+1} := \widetilde{Z}^h_i - \frac{t}{k}\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)(\widetilde{T}^l_i + v^1\widetilde{X}^l_i + v^2\widetilde{Y}^l_i + v^3\widetilde{Z}^l_i)\widetilde{Z}^j_i \tag{113}$$

for all $i \in \{0, \ldots, k-1\}$. The purpose of the defining equations listed in (113) is to guarantee that, as k tends to ∞, the points $\widetilde{\mathbf{q}}_0, \ldots, \widetilde{\mathbf{q}}_k$ form an increasingly accurate approximation of a curve $\mathbf{r}(t) = (\widetilde{q}^h(t))$ that satisfies the "geodesic" equation

$$\frac{d^2\widetilde{q}^h}{dt^2} + \Gamma^h_{jl}\frac{d\widetilde{q}^j}{dt}\frac{d\widetilde{q}^l}{dt} = 0. \tag{114}$$

(Remember that we are not working within the conceptual framework of differential geometry as yet. So our use of the term "geodesic" is at this point merely symbolic.) Given that our main emphasis in this paper is on physical interpretation rather than complete mathematical rigor, we will be content to provide an intuitively convincing argument for the validity of the assertion above rather than a detailed proof (which would not be difficult but tedious). Since the defining equations listed in (113) essentially represent a second order version of Euler's method of approximation, it is permissible to conclude that the limit of $\widetilde{\mathbf{q}}_k$ exists as the number of subdivision points k along the fixed interval $[0, t]$ increases to ∞. Thus, setting

$$\mathbf{r}(t) := (\widetilde{q}^h(t)) := \lim_{k\to\infty} \widetilde{\mathbf{q}}_k, \tag{115}$$

we may infer that for all sufficiently large values of k we have

$$\left.\frac{d^2\widetilde{q}^h}{dt^2}\right|_{it/k} \approx \frac{\dfrac{\widetilde{q}^h((i+2)t/k) - \widetilde{q}^h((i+1)/k)}{t/k} - \dfrac{\widetilde{q}^h((i+1)t/k) - \widetilde{q}^h(it/k)}{t/k}}{t/k}$$
$$\approx \frac{\dfrac{\widetilde{q}^h_{i+2} - \widetilde{q}^h_{i+1}}{t/k} - \dfrac{\widetilde{q}^h_{i+1} - \widetilde{q}^h_i}{t/k}}{t/k}$$
$$\approx \frac{(\widetilde{T}^h_{i+1} + v^1\widetilde{X}^h_{i+1} + v^2\widetilde{Y}^h_{i+1} + v^3\widetilde{Z}^h_{i+1}) - (\widetilde{T}^h_i + v^1\widetilde{X}^h_i + v^2\widetilde{Y}^h_i + v^3\widetilde{Z}^h_i)}{t/k}$$

$$
\begin{aligned}
&= -\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)(\widetilde{T}^l_i + v^1\widetilde{X}^l_i + v^2\widetilde{Y}^l_i + v^3\widetilde{Z}^l_i)\widetilde{T}^j_i \\
&\quad -v^1\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)(\widetilde{T}^l_i + v^1\widetilde{X}^l_i + v^2\widetilde{Y}^l_i + v^3\widetilde{Z}^l_i)\widetilde{X}^j_i \\
&\quad -v^2\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)(\widetilde{T}^l_i + v^1\widetilde{X}^l_i + v^2\widetilde{Y}^l_i + v^3\widetilde{Z}^l_i)\widetilde{Y}^j_i \\
&\quad -v^3\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)(\widetilde{T}^l_i + v^1\widetilde{X}^l_i + v^2\widetilde{Y}^l_i + v^3\widetilde{Z}^l_i)\widetilde{Z}^j_i \\
&= -\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)\frac{\widetilde{q}^l_{i+1}-\widetilde{q}^l_i}{t/k}\widetilde{T}^j_i - v^1\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)\frac{\widetilde{q}^l_{i+1}-\widetilde{q}^l_i}{t/k}\widetilde{X}^j_i \\
&\quad -v^2\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)\frac{\widetilde{q}^l_{i+1}-\widetilde{q}^l_i}{t/k}\widetilde{Y}^j_i - v^3\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)\frac{\widetilde{q}^l_{i+1}-\widetilde{q}^l_i}{t/k}\widetilde{Z}^j_i \\
&= -\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)\frac{\widetilde{q}^l_{i+1}-\widetilde{q}^l_i}{t/k}(\widetilde{T}^j_i + v^1\widetilde{X}^j_i + v^2\widetilde{Y}^j_i + v^3\widetilde{Z}^j_i) \\
&= -\Gamma^h_{jl}(\widetilde{\mathbf{q}}_i)\left(\frac{\widetilde{q}^l_{i+1}-\widetilde{q}^l_i}{t/k}\right)\left(\frac{\widetilde{q}^j_{i+1}-\widetilde{q}^j_i}{t/k}\right) \\
&\approx -\Gamma^h_{jl}(\widetilde{\mathbf{q}}(it/k))\left.\frac{d\widetilde{q}^l}{dt}\right|_{it/k}\left.\frac{d\widetilde{q}^j}{dt}\right|_{it/k}.
\end{aligned} \tag{116}
$$

Consequently, the parameterization defined in (115) does indeed satisfy the "geodesic" equation (114).

To continue with the proof of (111), we pick a presumably large integer m and define

$$
n := km, \quad \Delta t := \frac{t}{n}, \tag{117}
$$

and

$$
\begin{aligned}
\mathbf{V}_i(m) &:= -(\Gamma^k_{jl}(\mathbf{q}_{im})(T^l_{im} + v^1X^l_{im} + v^2Y^l_{im} + v^3Z^l_{im}))^4_{k,j=1}, \\
\widetilde{\mathbf{V}}_i &:= -(\Gamma^k_{jl}(\widetilde{\mathbf{q}}_i)(\widetilde{T}^l_i + v^1\widetilde{X}^l_i + v^2\widetilde{Y}^l_i + v^3\widetilde{Z}^l_i))^4_{k,j=1}
\end{aligned} \tag{118}
$$

for all $i \in \{0, \ldots, k-1\}$. The initial segment of the random walk that we are now going to consider begins at $\mathbf{q}_0$ and ends at $\mathbf{q}_m$, and in general, the ith segment begins at $\mathbf{q}_{im}$ and ends at $\mathbf{q}_{(i+1)m}$ for $i \in \{0, \ldots, k-1\}$. In order to apply (98) to the ith segment, we replace $\mathbf{q}_0$ (in our derivation of (98)) by $\mathbf{q}_{im}$, $\mathbf{q}_n$ by $\mathbf{q}_{(i+1)m}$, t by t/k, n by m, $\mathbf{T}_0$ by $\mathbf{T}_{im}$, $\mathbf{X}_0$ by $\mathbf{X}_{im}$, $\mathbf{Y}_0$ by $\mathbf{Y}_{im}$, $\mathbf{Z}_0$ by $\mathbf{Z}_{im}$, and $\mathbf{r}$, as defined in (94), by

$$
\mathbf{s}_i(m) := \mathbf{q}_{im} + \frac{t}{k}(\mathbf{T}_{im} + v^1\mathbf{X}_{im} + v^2\mathbf{Y}_{im} + v^3\mathbf{Z}_{im}). \tag{119}
$$

Given these substitutions, (98) and (112) imply that

$$\begin{aligned}&\lim_{m\to\infty} P\left(\|\mathbf{q}_{(i+1)m} - \mathbf{s}_i(m)\| \leqslant \delta/k\right)\\ &\geqslant \lim_{m\to\infty} P\left(\|\mathbf{q}_{(i+1)m} - \mathbf{s}_i(m)\| \leqslant \nu\lambda(t/k)^2 e^{\lambda t/k} + 3\nu\varepsilon t/k\right) = 1\end{aligned} \tag{120}$$

for all $i \in \{0, \ldots, k-1\}$. Setting

$$\begin{aligned}a_i(m) &:= \|\mathbf{q}_{im} - \widetilde{\mathbf{q}}_i\|,\\ b_i(m) &:= \|\mathbf{T}_{im} - \widetilde{\mathbf{T}}_i\| + \|\mathbf{X}_{im} - \widetilde{\mathbf{X}}_i\| + \|\mathbf{Y}_{im} - \widetilde{\mathbf{Y}}_i\| + \|\mathbf{Z}_{im} - \widetilde{\mathbf{Z}}_i\|,\end{aligned} \tag{121}$$

we may apply the definitions of $\widetilde{\mathbf{q}}_i$ and $\mathbf{s}_i(m)$ to infer that

$$\begin{aligned}a_{i+1}(m) &\leqslant \|\mathbf{q}_{(i+1)m} - \mathbf{s}_i(m)\| + \|\mathbf{s}_i(m) - \widetilde{\mathbf{q}}_i - (\widetilde{\mathbf{q}}_{i+1} - \widetilde{\mathbf{q}}_i)\|\\ &\leqslant \|\mathbf{q}_{(i+1)m} - \mathbf{s}_i(m)\| + \|\mathbf{q}_{im} - \widetilde{\mathbf{q}}_i\|\\ &\quad + \frac{t}{k}(\|\mathbf{T}_{im} - \widetilde{\mathbf{T}}_i\| + \|\mathbf{X}_{im} - \widetilde{\mathbf{X}}_i\| + \|\mathbf{Y}_{im} - \widetilde{\mathbf{Y}}_i\| + \|\mathbf{Z}_{im} - \widetilde{\mathbf{Z}}_i\|)\\ &= \|\mathbf{q}_{(i+1)m} - \mathbf{s}_i(m)\| + a_i(m) + \frac{t}{k} b_i(m).\end{aligned} \tag{122}$$

Moreover, according to (109) and (112) (with $m = t/(k\Delta t)$, $\mathbf{T}_{(i+1)m}$, $\mathbf{T}_{im}$, and $\mathbf{V}_i(m)$ in place of n, $\mathbf{T}_n$, $\mathbf{T}_0$, and $\mathbf{V}_0$, respectively), we have

$$\begin{aligned}\|\mathbf{T}_{(i+1)m} - \widetilde{\mathbf{T}}_{i+1}\| &\leqslant \|\mathbf{T}_{(i+1)m} - (\mathbf{T}_{im} + m\mathbf{V}_i(m)\mathbf{T}_{im}\Delta t)\|\\ &\quad + \|\mathbf{T}_{im} - \widetilde{\mathbf{T}}_i\| + \|m\mathbf{V}_i(m)\mathbf{T}_{im}\Delta t - (\widetilde{\mathbf{T}}_{i+1} - \widetilde{\mathbf{T}}_i)\|\\ &\leqslant \frac{\delta}{4k} + \frac{4t\nu^2 D}{k}(|\bar{x}_i - \mu_x| + |\bar{y}_i - \mu_y| + |\bar{z}_i - \mu_z|)\\ &\quad + \|\mathbf{T}_{im} - \widetilde{\mathbf{T}}_i\| + \frac{t}{k}\|\mathbf{V}_i(m)\mathbf{T}_{im} - \widetilde{\mathbf{V}}_i\widetilde{\mathbf{T}}_i\|,\end{aligned} \tag{123}$$

where the sample means $\bar{x}_i$, $\bar{y}_i$ and $\bar{z}_i$ are taken over indices ranging from im to $(i+1)m-1$. Given the definitions of $\widetilde{\mathbf{V}}_i$ and $\mathbf{V}_i(m)$ in (118), the inequalities (105) and (106) allow us to conclude that

$$\begin{aligned}&\|\mathbf{V}_i(m)\mathbf{T}_{im} - \widetilde{\mathbf{V}}_i\widetilde{\mathbf{T}}_i\|\\ &\quad\leqslant \|(\Gamma^k_{jl}(\mathbf{q}_{im}) - \Gamma^k_{jl}(\widetilde{\mathbf{q}}_i))(T^l_{im} + v^1 X^l_{im} + v^2 Y^l_{im} + v^3 Z^l_{im})T^j_{im}\|\\ &\qquad + \|\Gamma^k_{jl}(\widetilde{\mathbf{q}}_i)(T^l_{im} + v^1 X^l_{im} + v^2 Y^l_{im} + v^3 Z^l_{im})(T^j_{im} - \widetilde{T}^j_i)\|\\ &\qquad + \|\Gamma^k_{jl}(\widetilde{\mathbf{q}}_i)(T^l_{im} - \widetilde{T}_i + v^1(X^l_{im} - \widetilde{X}^l_i) + v^2(Y^l_{im} - \widetilde{Y}^l_i) + v^3(Z^l_{im} - \widetilde{Z}^l_i))\widetilde{T}^j_i\|\\ &\quad\leqslant 4\nu^2 D\|\mathbf{q}_{im} - \widetilde{\mathbf{q}}_i\| + 4\nu D\|\mathbf{T}_{im} - \widetilde{\mathbf{T}}_i\|\\ &\qquad + 4\nu D(\|\mathbf{T}_{im} - \widetilde{\mathbf{T}}_i\| + \|\mathbf{X}_{im} - \widetilde{\mathbf{X}}_i\| + \|\mathbf{Y}_{im} - \widetilde{\mathbf{Y}}_i\| + \|\mathbf{Z}_{im} - \widetilde{\mathbf{Z}}_i\|)\\ &\quad\leqslant 4\nu^2 D a_i(m) + 8\nu D b_i(m).\end{aligned} \tag{124}$$

Hence

$$\|\mathbf{T}_{(i+1)m} - \widetilde{\mathbf{T}}_{i+1}\| \leqslant \frac{\delta}{4k} + \frac{4tv^2D}{k}(|\bar{x}_i - \mu_x| + |\bar{y}_i - \mu_y| + |\bar{z}_i - \mu_z|) + \|\mathbf{T}_{im} - \widetilde{\mathbf{T}}_i\| + \frac{t}{k}(4v^2Da_i(m) + 8vDb_i(m)). \quad (125)$$

Using the inequalities in (110), it can be shown that a completely analogous estimate is valid for $\|\mathbf{X}_{(i+1)m} - \widetilde{\mathbf{X}}_{i+1}\|$ as well, i.e.,

$$\|\mathbf{X}_{(i+1)m} - \widetilde{\mathbf{X}}_{i+1}\| \leqslant \frac{\delta}{4k} + \frac{4tv^2D}{k}(|\bar{x}_i - \mu_x| + |\bar{y}_i - \mu_y| + |\bar{z}_i - \mu_z|) + \|\mathbf{X}_{im} - \widetilde{\mathbf{X}}_i\| + \frac{t}{k}(4v^2Da_i(m) + 8vDb_i(m)), \quad (126)$$

and similarly for $\|\mathbf{Y}_{(i+1)m} - \widetilde{\mathbf{Y}}_{i+1}\|$ and $\|\mathbf{Z}_{(i+1)m} - \widetilde{\mathbf{Z}}_{i+1}\|$. Adding up these four estimates yields

$$b_{i+1}(m) \leqslant \frac{\delta}{k} + \frac{\alpha}{k}(|\bar{x}_i - \mu_x| + |\bar{y}_i - \mu_y| + |\bar{z}_i - \mu_z|) + \left(1 + \frac{\alpha}{k}\right) b_i(m) + \frac{\alpha}{k} a_i(m) \quad (127)$$

where

$$\alpha := \max\{16tv^2D, 32tvD, 2, t\}. \quad (128)$$

(Note: the last two elements—2 and t—in the set above are not needed for (127), but they will be useful in deriving (130) and (131) below.) Using Chebyshev's inequality, as we did earlier, it is easy to see that

$$\lim_{m\to\infty} P(\alpha(|\bar{x}_i - \mu_x| + |\bar{y}_i - \mu_y| + |\bar{z}_i - \mu_z|) \leqslant \delta) = 1 \quad (129)$$

and therefore,

$$\lim_{m\to\infty} P\left(b_{i+1}(m) \leqslant \frac{\delta\alpha}{k} + \left(1 + \frac{\alpha}{k}\right) b_i(m) + \frac{\alpha}{k} a_i(m)\right) \geqslant \lim_{m\to\infty} P\left(b_{i+1}(m) \leqslant \frac{2\delta}{k} + \left(1 + \frac{\alpha}{k}\right) b_i(m) + \frac{\alpha}{k} a_i(m)\right) = 1 \quad (130)$$

for all $i \in \{0, \ldots, k-1\}$. Moreover, according to (120) and (122), we also have

$$\lim_{m\to\infty} P\left(a_{i+1}(m) \leqslant \frac{\delta\alpha}{k} + \left(1+\frac{\alpha}{k}\right)a_i(m) + \frac{\alpha}{k}b_i(m)\right) \geqslant \lim_{m\to\infty} P\left(a_{i+1}(m) \leqslant \frac{\delta}{k} + a_i(m) + \frac{t}{k}b_i(m)\right) = 1 \tag{131}$$

for all $i \in \{0, \ldots, k-1\}$. To proceed, we will use (130) and (131) to prove by induction that

$$1 = \lim_{m\to\infty} P\left(a_i(m) \leqslant \frac{\delta((1+2\alpha/k)^i - 1)}{2}\right),$$
$$1 = \lim_{m\to\infty} P\left(b_i(m) \leqslant \frac{\delta((1+2\alpha/k)^i - 1)}{2}\right). \tag{132}$$

for all $i \in \{0, \ldots, k\}$. Since $a_0(m) = b_0(m) = 0$, the equalities above are obviously valid for $i = 0$, and if they are valid for some index $i \in \{0, \ldots, k-1\}$, then, according to (131), we have

$$\begin{aligned}
&\lim_{m\to\infty} P\left(a_{i+1}(m) \leqslant \frac{\delta((1+2\alpha/k)^{i+1}-1)}{2}\right)\\
&\geqslant \lim_{m\to\infty} P\left(\frac{\delta\alpha}{k} + \left(1+\frac{\alpha}{k}\right)a_i(m) + \frac{\alpha}{k}b_i(m) \leqslant \frac{\delta((1+2\alpha/k)^{i+1}-1)}{2}\right)\\
&\geqslant \lim_{m\to\infty} P\left(\frac{\delta\alpha}{k} + \left(1+\frac{2\alpha}{k}\right)\frac{\delta((1+2\alpha/k)^{i}-1)}{2} \leqslant \frac{\delta((1+2\alpha/k)^{i+1}-1)}{2}\right)\\
&= \lim_{m\to\infty} P\left(\frac{\delta((1+2\alpha/k)^{i+1}-1)}{2} \leqslant \frac{\delta((1+2\alpha/k)^{i+1}-1)}{2}\right) = 1,
\end{aligned} \tag{133}$$

as desired. Using (130) instead of (131), the proof with $b_{i+1}(m)$ in place of $a_{i+1}(m)$ is easily seen to be completely analogous. Consequently, the equalities in (132) are indeed valid for all $i \in \{0, \ldots, k\}$. Setting i equal to k yields

$$\lim_{m\to\infty} P\left(a_k(m) \leqslant \frac{\delta((1+2\alpha/k)^k - 1)}{2}\right) = 1. \tag{134}$$

Since

$$\frac{(1+2\alpha/k)^k - 1}{2} \leqslant \alpha e^{2\alpha} \tag{135}$$

for all $k \geqslant 1$, it follows that

$$\lim_{m\to\infty} P(\|\mathbf{q}_{km} - \widetilde{\mathbf{q}}_k\| \leqslant \delta\alpha e^{2\alpha}) = 1 \tag{136}$$

for all k that are large enough to satisfy the conditions in (112). Furthermore, according to (115), it is also true that for all sufficiently large k we have

$$\|\widetilde{\mathbf{q}}_k - \mathbf{r}(t)\| \leqslant \delta\alpha e^{2\alpha}, \tag{137}$$

and therefore,

$$\begin{aligned} &\lim_{m\to\infty} P(\|\mathbf{q}_{mk} - \mathbf{r}(t)\| \leqslant 2\delta\alpha e^{2\alpha}) \\ &\quad\geqslant \lim_{m\to\infty} P(\|\mathbf{q}_{mk} - \widetilde{\mathbf{q}}_k\| + \|\widetilde{\mathbf{q}}_k - \mathbf{r}(t)\| \leqslant 2\delta\alpha e^{2\alpha}) \\ &\quad\geqslant \lim_{m\to\infty} P(\|\mathbf{q}_{mk} - \widetilde{\mathbf{q}}_k\| \leqslant \delta\alpha e^{2\alpha}) = 1. \end{aligned} \tag{138}$$

Since δ was initially chosen to be an arbitrary positive number, this result implies that

$$\lim_{n\to\infty} P(\|\mathbf{q}_n - \mathbf{r}(t)\| \leqslant \delta) = \lim_{m\to\infty} P(\|\mathbf{q}_{mk} - \mathbf{r}(t)\| \leqslant \delta) = 1 \tag{139}$$

for all $\delta > 0$. Thus the proof of (111) is complete.

*Remark***.** The assertion of equality in the concluding equation above with regard to the limits of $P(\|\mathbf{q}_n - \mathbf{r}(t)\| \leqslant \delta)$ and $P(\|\mathbf{q}_{mk} - \mathbf{r}(t)\| \leqslant \delta)$ as n and m tend to ∞ hardly needs any justification, but perhaps we should mention that for large values of k the positions of $\mathbf{q}_{(k-1)m}$, $\mathbf{q}_n$, and $\mathbf{q}_{km}$ are almost identical whenever $(k-1)m \leqslant n \leqslant km$ because the lengths $\|\mathbf{T}_i\|$, $\|\mathbf{X}_i\|$, $\|\mathbf{Y}_i\|$, $\|\mathbf{Z}_i\|$ are universally bounded. Consequently, if the positions of $\mathbf{q}_{(k-1)m} \approx \mathbf{r}((k-1)t/k)$ and $\mathbf{q}_{km} \approx \mathbf{r}(t)$ are statistically stable, then the same must be true for the position of $\mathbf{q}_n$.

In order to put the preceding discussion in a proper perspective, we now assume that the functions Γ^k_{jl} are the Christoffel symbols associated with a space–time metric tensor $\mathbf{G} = (g_{jl})$ defined on M (or $\mathbb{R}^4$). Then, at any point $\mathbf{q}_0 \in M$ we can find a set of basis vectors $\{\mathbf{T}_0, \mathbf{X}_0, \mathbf{Y}_0, \mathbf{Z}_0\}$ in the tangent space at $\mathbf{q}_0$ such that the Minkowski inner product in that particular tangent space satisfies the equation

$$(t\mathbf{T}_0 + x\mathbf{X}_0 + y\mathbf{Y}_0 + z\mathbf{Z}_0)^t\mathbf{G}(t\mathbf{T}_0 + x\mathbf{X}_0 + y\mathbf{Y}_0 + z\mathbf{Z}_0) = t^2 - x^2 - y^2 - z^2 \tag{140}$$

for all $t, x, y, z \in \mathbb{R}$. Given this setup, the arguments presented above allow us to infer that, as Δt approaches zero, a random walk starting at $\mathbf{q}_0$ contracts with probability 1 to a geodesic $\mathbf{r}(t)$ in space–time that satisfies the initial conditions $\mathbf{r}(0) = \mathbf{q}_0$ and $\mathbf{r}'(0) = \mathbf{T}_0 + v^1\mathbf{X}_0 + v^2\mathbf{Y}_0 + v^3\mathbf{Z}_0$ (the latter equality follows from the observation that for all sufficiently large values of k we have $\mathbf{r}'(0) \approx (\widetilde{\mathbf{q}}_1 - \widetilde{\mathbf{q}}_0)/(t/k) = \widetilde{\mathbf{T}}_0 + v^1\widetilde{\mathbf{X}}_0 + v^2\widetilde{\mathbf{Y}}_0 + v^3\widetilde{\mathbf{Z}}_0 = \mathbf{T}_0 + v^1\mathbf{X}_0 + v^2\mathbf{Y}_0 + v^3\mathbf{Z}_0$).

*Remark***.** It is understood that the tangent space to the q^k-coordinate system (that is, to $\mathbb{R}^4$) at $\mathbf{q}_0$ is here identified with $\mathbb{R}^4$ itself. Such an identification makes no sense within the conceptual framework of differential geometry, but it is perfectly legal as far as our description of random walks is concerned because, in essence, our proof of (111) only depended on the topological structure induced by the Euclidean norm on $\mathbb{R}^4$ rather than the differential structure of the space–time manifold that $\mathbf{G}$ is associated with.

In order to show that even in a general relativistic setting proper time can be measured in terms of statistical spread, we set

$$\begin{aligned} S_x^2 &:= \frac{1}{n-1}\sum_{i=0}^{n-1}(x_i - \bar{x})^2, \\ S_y^2 &:= \frac{1}{n-1}\sum_{i=0}^{n-1}(y_i - \bar{y})^2, \\ S_z^2 &:= \frac{1}{n-1}\sum_{i=0}^{n-1}(z_i - \bar{z})^2. \end{aligned} \tag{141}$$

Given these definitions, a good approximate measure of the proper time corresponding to the first n steps of a random walk starting at $\mathbf{q}_0$ is the quantity

$$n\Delta t\sqrt{S_x^2 + S_y^2 + S_z^2}. \tag{142}$$

To understand why this is so, we observe that the expected value of the sum $S_x^2 + S_y^2 + S_z^2$ is

$$\begin{aligned} E(S_x^2 + S_y^2 + S_z^2) &= \mathrm{Var}(x_i) + \mathrm{Var}(y_i) + \mathrm{Var}(y_i) \\ &= E(x_i^2 + y_i^2 + z_i^2) - E(x_i)^2 - E(y_i)^2 - E(z_i)^2 \\ &= 1 - \mu_x^2 - \mu_y^2 - \mu_z^2 \\ &= 1 - \|\mathbf{v}\|^2. \end{aligned} \tag{143}$$

Consequently, for large values of n, the expression in (142) will be approximately equal to

$$n\Delta t\sqrt{1-\|\mathbf{v}\|^2} = t\sqrt{1-\|\mathbf{v}\|^2}. \tag{144}$$

This result is in perfect agreement with the prediction of the standard continuous model, because along any geodesic the Minkowski inner product of the derivative vector is known to remain constant, and, according to (140), the proper time over the interval $[0, t]$ is therefore equal to

$$\begin{aligned}\int_0^t \sqrt{\mathbf{r}'(\tau)^t\mathbf{G}(\mathbf{r}(\tau))\mathbf{r}'(\tau)}\,d\tau &= t\sqrt{\mathbf{r}'(0)^t\mathbf{G}(\mathbf{r}(0))\mathbf{r}'(0)}\\ &= t\sqrt{(\mathbf{T}_0+v^1\mathbf{X}_0+v^2\mathbf{Y}_0+v^3\mathbf{Z}_0)^t\mathbf{G}(\mathbf{q}_0)(\mathbf{T}_0+v^1\mathbf{X}_0+v^2\mathbf{Y}_0+v^3\mathbf{Z}_0)}\\ &= t\sqrt{1-\|\mathbf{v}\|^2},\end{aligned} \tag{145}$$

as desired. Thus it appears that the notions of proper time and randomness are indeed equivalent.

4. A GENUINELY PROBABILISTIC MODEL?

What we accomplished in Sec. 3 was to demonstrate that the completely coherent world of general relativity emerges from a discrete model of motion in the statistical limit. What we did not establish is the coherence of a genuinely probabilistic world in itself. In particular, we did not show how the expected position of the endpoint $\mathbf{q}_n$ of a random walk can be computed or how recordings of such expected positions in different coordinate systems can be harmonized. It is true that for small values of Δt different observers will locate the endpoint of a random walk with high probability in close proximity to the same geodesic, but the actual expected positions may still differ slightly.

To create a genuinely discrete model of general relativity in the case where a positive lower bound for Δt (such as the Planck length) does exist may require a more radical departure from spatial and temporal continuity than the one proposed in the present paper. However, if we allow the values of Δt to be arbitrarily small, then a consistent probabilistic model can be constructed if we take an approach similar to Feynman's path integral formulation of quantum mechanics whereby the motion of a particle between two points is analyzed with reference to the totality of all paths that connect these two points. To see this, we only need to identify the motion along a geodesic $\mathbf{r}(t)$ over an interval $[0, t]$ with an infinite sequence of random walks $\mathcal{W}_n = (\mathbf{q}_0(n), \ldots, \mathbf{q}_{2^n}(n))$ each of which has its own step length $\Delta t_n := t/2^n$. If $\mathbf{q}_0 = \mathbf{q}_0(n)$ is the common starting point of these random walks, $\mathbf{v}$ the common "velocity vector," and $\{\mathbf{T}_0, \mathbf{X}_0, \mathbf{Y}_0, \mathbf{Z}_0\}$ the common set of initial basis vectors, then, as n tends to ∞, the sequence $\mathbf{q}_{2^{n-m}k}(n)$ converges with probability 1 to $\mathbf{r}(kt/2^m)$ for all $m \in \mathbb{N}$ and all $k \in \{0, \ldots, 2^m\}$. In other words, the geodesic $\mathbf{r}(t)$ is uniquely determined by the sequence $(\mathcal{W}_n)_{n=1}^{\infty}$ on the dense set $\{kt/2^m | m \in \mathbb{N},\ 0 \leqslant k \leqslant 2^m\} \subset [0, t]$, and by continuous extension it is uniquely determined on the entire interval $[0, t]$. It is in this sense then that we may consider a probabilistic model of motion to be equivalent to the standard continuous model even in a general relativistic setting.

ACKNOWLEDGMENTS

I would like to thank my former colleague Gaston Griggs for many stimulating discussions concerning the nature of time and, in particular, for his suggestion that physical reality is essentially nontemporal and is best thought of as an ordered sequence of discrete static images.

REFERENCE

1. J. Barbour, *The End of Time* (Oxford University Press, Oxford, 2000).