

Article

# Visit Probability in Space–Time Prisms Based on Binomial Random Walk

Deepak Elias  and Bart Kuijpers \* 

Databases and Theoretical Computer Science Research Group, Data Science Institute, UHasselt—Hasselt University, Agoralaan, Gebouw D, 3590 Diepenbeek, Belgium

\* Correspondence: bart.kuijpers@uhasselt.be

Received: 15 July 2020; Accepted: 13 September 2020; Published: 18 September 2020



**Abstract:** Space–time prisms are used to model the uncertainty of space–time locations of moving objects between (for instance, GPS-measured) sample points. However, not all space–time points in a prism are equally likely and we propose a simple, formal model for the so-called “visit probability” of space–time points within prisms. The proposed mathematical framework is based on a binomial random walk within one- and two-dimensional space–time prisms. Without making any assumptions on the random walks (we do not impose any distribution nor introduce any bias towards the second anchor point), we arrive at the conclusion that binomial random walk-based visit probability in space–time prisms corresponds to a hypergeometric distribution.

**Keywords:** geographic information science; time geography; space–time prisms; random walk; uncertainty; probability; visit probability; binomial distribution; hypergeometric distribution

## 1. Introduction

Due to the discrete nature of moving object data, typically measured by GPS devices at distinct moments in time, there exists an inherent uncertainty between measured locations. Modelling the aforementioned uncertainty is necessary for a wide range of applications and space–time prisms have been extensively investigated as a tool for this purpose. They have been used in studying accessibility [1–3], human behaviour [4–6], criminal offense patterns [7], as well as for the study of transportation [8–11]. A space–time prism is the demarcation of all the possible space–time locations accessible by a moving object between two measured space–time locations (called anchor points), given a physical constraint on its velocity. The spatial projection of the prism, also called the potential path area, is an envelope of the spatial whereabouts of the moving object between the measured spatial locations. Further discussions and mathematical formulations for the various components of time geography are detailed by Miller [12].

In its basic form, a space–time prism lacks an internal structure, meaning that the space–time points in a prism are not distinguished and are assigned equal importance. Since there is an infinite number of velocity bound paths within the prism, we can observe that each point inside a prism is visited by infinitely many of them (with the exception of some boundary points of the prism). This observation leads to the conclusion that there is no a priori reason to distinguish between space–time points inside the prism, unless some further assumption is adopted besides the velocity bound. On the other hand, some basic trajectory reconstruction methods, the first being linear interpolation between anchor points [13], assumes that the linear path between the anchor points is the most likely one among the infinite collection of paths bounded by the prism. Still, in many applications, such as animal or human movement [14–16], it is obvious and plausible that certain points in the prism, such as those on a linear interpolation path, should be considered more likely than points towards the boundary of the prism that require a considerable detour to visit.

For many applications, it is also desirable to know the likelihood of points in a prism to be visited. We can take the example of bird migration data from animal ecology [17,18]. Systems like Argos [19] can only collect the measured data points sparsely to save battery life. While constructing a prism from such sparse data points might result in a large potential path area, we will be able to answer certain queries that are prominent in animal ecology. For instance, if the location of a bird colony with a disease like influenza is known, a relevant query could be the probability of a migrating bird having been in contact with the diseased bird colony. Other queries in this application concern the probability of a bird having visited some food sources. Another application is related to the alibi query, where it is relevant to know with which likelihood a potential suspect could have visited a crime scene or could have met another suspect [20]. Obviously, the use of technology not only contributes to pursuing the safety of society. Over-policing using technology in the urban environment has been criticized too, and raises, for instance, privacy issues. For this type of concerns, we refer to [21].

In the literature, several proposals have been made to assign probability values to space–time points within a prism, thus providing the prism with an internal structure that expresses the unequal movement opportunities within the prism. A great deal of progress has been achieved by leveraging parametric probability distributions as models for the expected interior structure of space–time prisms.

Whereas in the basic model, the value assigned to space–time points is Boolean (possible or impossible to visit), an improved accessibility metric would enable a better understanding of various paths and movement opportunities available for the moving object within the prism [22–26].

The notion of probability distributions in space–time prisms for movement in the plane, has become known as “visit probability” and has been studied by various researchers. Winter and Yin [24] model visit probability in a prism using the theory of random walks and they arrive at the result that the internal distribution is of a bivariate multinomial nature at any given time. Their work was initially based on undirected random walks, which was extended later to include the directional aspect by introducing a bias to the destination anchor in their model. The distribution they obtained from the undirected model was then translated to establish that the distribution is at its peak around the axis connecting the first and the second anchor point, that is, at the shortest path. Further, they extend the model from the discrete to continuous space by applying the continuous analog of the bivariate multinomial distribution, namely the bivariate normal distribution. In order to maintain the constraints on movement as determined by the prism parameters, they adjust the standard deviation of the distribution to fit the directional bias towards the second anchor point. They also clip the distribution along the prism boundaries.

The work of Winter and Yin is extended by Song and Miller [26] with separate methods to simulate the visit probability in prisms for both the discrete and continuous case. Like Winter and Yin, they used a random walk model for the discrete case, but they modify it to include various other parameters such as the directional element (to account for the second anchor point), the maximum velocity and the remaining time budget (by actively revising the random walk matrix after each time step). In the continuous case, they model a framework on the Brownian bridge approach. Brownian motion is a continuous stochastic process which was initially designed to model the movement of infinitely-often colliding atomic particles. Whereas, Brownian bridges are a variant of Brownian motion used to model movement between two locations, and are used to evaluate stock prices as well as patterns of migration in the animal ecosystem [27]. As Brownian bridges in two dimensions effectively produce a probability distribution of a bivariate normal nature [28], Song and Miller conclude visit probability to be of the same nature. While the result reinforces the work done by Winter and Yin, it is not based on many of the assumptions made in the earlier work. Song and Miller base their work on an established movement theory. In order to ensure that the Brownian bridges stay within the limits of the prism boundary, they use the notion of truncated distributions, a conceptual method to restrict the domain of a parent distribution [29]. They use a variable called the “dispersion parameter” to capture the variations in the movement ability of the object within the prism. The value of the dispersion parameter can be based on real-life movement data. Song and Miller’s simulations of visit probability produce

results similar to the intuitive expectations where the distribution is the highest along the axis of the prism. Some further approaches to visit probability, less relevant to the present paper, are discussed in the related work section.

This paper explores the approaches started in Winter and Yin [24] and in Song and Miller [26] in more detail. We propose a mathematical framework for random walk-based visit probability within one- and two-dimensional space–time prisms. In fact, like these authors, we consider movement in a prism that is constrained to what Burns calls fine-grid networks (see pages 33–34 of [30]). Our random walk, at each point, has the (random) choice between two directions: going left on the grid or going right on the grid. However, unlike these authors, we make no assumptions on the random walks, we impose no distributions, we have no truncations and we do not introduce a bias towards the second anchor point. Simply by defining the random walk on a grid and making no further assumptions, we arrive at a more natural conclusion: random walk-based visit probability in space–time prisms corresponds to a hypergeometric distribution. This conclusion reflects what is really going on in this situation, when no distributions are a priori imposed on the walks in the prism. We then further explore this hypergeometric distribution and investigate, for instance, what happens when the grid is refined. This is relevant to the problem of how to extend this distribution from the grid network to the complete prism. For the complete one-dimensional, we can obtain a distribution by extending the result for grids to the complete prism by using barycentric coordinates. This results in a linear interpolation of the grid case. This approach gives a more continuous result, as opposed to the voxel-based approach of, for example, References [31,32] and of the random-walk part of [24,26]. In the one-dimensional case, fine-grid networks naturally fill the space–time prism. This is not the case for the two-dimensional case, where the fine-grid network occupies a volume of a pyramidal-shaped prism within the classical space–time prism. We conclude the paper by exploring example applications where our approach could be usefully applied. These applications include movement modelling in space–time space applied to human and animal movement and interaction, trajectory data cleaning and analysis and trajectory data querying.

**Organisation:** This paper is organized as follows. After some introductory definitions in Section 3, we define and explore our random walk model for one-dimensional prisms in Section 4. In Section 5, we deal with the case of movement in a two-dimensional space. We discuss example application areas in Section 6. Our conclusion is given in Section 7.

## 2. Related Work

The notion of space–time prisms was introduced in the early 1970s in the area of “time geography” by Hägerstrand [33] to model the accessibility of locations, also with respect to time. In the 1990s, the space–time prism model entered into the GIS community with works by Janelle and Goodchild and Miller [34,35], where it was mainly used as a method for analyzing the potential activity of people in some environment (a well-known reference in this context is [35]). Later on, this model became popular in the area of spatio-temporal and moving object databases due to the work by Pfoser and Jensen [36]. In this field, it was further studied by Egenhofer and Hornsby [37,38].

Classical space–time prisms, as introduced by Hägerstrand, assume a local, albeit uniform travel velocity in an isotropic space such as the two-dimensional plane. This assumption ignores the fact that the travel possibilities of an individual are often confined to some type of transportation network which has link-dependent travel velocities. This insight has led to the study of space–time prisms on transportation networks [4,8,35,39].

Many analytical and computational tools for describing prisms in planar space and on transportation networks have been developed [4,6,12,35,40]. Combined with a strong theoretical foundation [9,41], these tools have led to a wide range of applications that use space–time prisms on their vast amounts of trajectory (sample) data. These applications include transportation planning [42], social research and crime analysis [7,43] and also epidemiological and environmental risk assessment [44–46].

In many contexts, prisms are both used to measure an individual's accessibility to a certain environment, as well as to model the uncertainty about the location of moving objects between known locations. Intersections between prisms express the potential for human contact [47] and their disjointness can be used as an alibi, for instance, in a criminal investigation [20,37].

Concerning the topic of assigning probabilities to space–time points in a prism, we mention the following works. An approach is to model the probable movement space using various methods to available empirical data, such as the “surface generation methods”. Xie and Yan [22] describe the application of the “kernel density estimation” method to model estimation of traffic accidents in networked space. Downs [25] extends the approach to introduce kernel density estimation to time–geography by adding the space–time prism constraints to the model, which she named the “time–geographic density estimation” method. All sampled events for a kernel density estimation are selected randomly or periodically from a set of sequenced control points in between the origin and the destination. The maximum velocity and time required for travel are used to calculate the bandwidth of the kernel density estimation, while the potential path area as defined by the origin and the destination replaces the kernel function. The method can be used to derive a probability metric for the moving objects spatial locations at various times in planar space–time prisms. Downs et al. [32] further propose the “Inverse Distance Weighting” to realize the interior of prisms. They introduce the “Voxel-based” probabilistic space–time prisms for animal movement and habitat use analysis in their work.

Long et al. [48] use skew-normal distributions as a heuristic to model movement probabilities within space–time prisms, applying their method to analyze wildlife, cyclist and athlete movement data. Later on, Long [49] compares methods for kinematic interpolation of movement data based on Bézier and Catmull–Rom curves with classic interpolation methods for correlated random walks, caribou, cyclist, hurricane, and athlete tracking data. These authors conclude that kinematic methods are superior for fast-moving objects. More recently, studies on the internal structure of space–time prisms have considered behavioral or contextual information [15]. Loraamm further expands on this by considering behavioral observations as the primary parameter for the study of movement in her work on “Agent-Based Simulation” [50].

In broad terms, these approaches to introduce probability in prisms can be classified into two categories: the “quantitative” approach and the “model” approach. The quantitative approach is data-driven and uses empirical or historical data to build a visiting likelihood, often using some machine learning or data mining techniques. In this paper, we follow the model approach in which a mathematical framework or model is built, based on some model of or approach to movement.

### 3. Definitions and Preliminaries

We denote the underlying space in which we consider the movement of objects to take place by  $\mathbf{S}$ . In this paper,  $\mathbf{S}$  is taken to be the real line  $\mathbf{R}$  or the real plane  $\mathbf{R}^2$ . The spatio-temporal space in which movement takes place is then denoted by  $\mathbf{S} \times \mathbf{R}$ , in which  $\mathbf{R}$  is the temporal component. We use  $p, q, r, \dots$  to refer to spatial points (in  $\mathbf{S}$ ) and  $t$  to refer to the time coordinate (always, with or without indices). We use  $x, y, \dots$  to refer to spatial coordinates of spatial points. For example, for a space–time point  $(p, t) \in \mathbf{R}^2 \times \mathbf{R}$ , we have  $(p, t) = (x, y, t)$  if  $(x, y)$  are the spatial coordinates of  $p$  in  $\mathbf{R}^2$ . We assume that a distance function  $d$  is defined on  $\mathbf{S}$ , which we take to be the standard Euclidean distance function for  $\mathbf{R}$  and  $\mathbf{R}^2$ . A natural extension of movement on the line  $\mathbf{R}$  is movement on a transportation (or road) networks in  $\mathbf{R}^2$ , in which case, the distance function can be derived from the shortest-path distance on road networks [8]. We are now ready to define the notion of space–time prism for movement in the space  $\mathbf{S}$ .

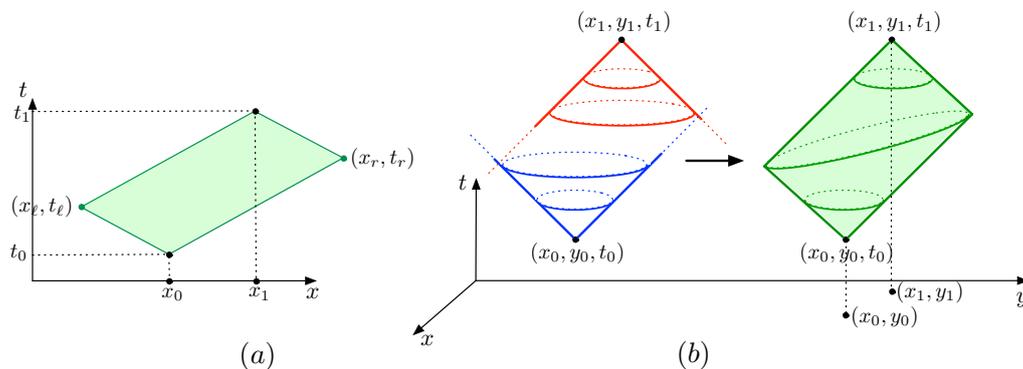
**Definition 1.** Let  $(p_0, t_0), (p_1, t_1) \in \mathbf{S} \times \mathbf{R}$  be two space–time points with  $t_0 < t_1$  and let  $v_{max} \in \mathbf{R}$  with  $v_{max} \geq 0$  be a bound on the speed of a moving object. The space–time prism  $\mathcal{P}(p_0, t_0, p_1, t_1, v_{max})$  is defined to be the set of all space–time points  $(p, t)$  that are in the intersection of the future cone (or bottom cone)

$C^-(p_0, t_0, v_{max}) := \{(p, t) \in \mathbf{S} \times \mathbf{R} \mid d(p_0, p) \leq (t - t_0) \cdot v_{max} \wedge t_0 \leq t\}$  and the past cone (or top cone)  
 $C^+(p_1, t_1, v_{max}) := \{(p, t) \in \mathbf{S} \times \mathbf{R} \mid d(p_1, p) \leq (t_1 - t) \cdot v_{max} \wedge t \leq t_1\}$ .

The space–time points  $(p_0, t_0)$  and  $(p_1, t_1)$  are called the anchor points of the space–time prism  $\mathcal{P}(p_0, t_0, p_1, t_1, v_{max})$ .

For example, for movement in the plane  $\mathbf{R}^2$ , the prism  $\mathcal{P}(p_0, t_0, p_1, t_1, v_{max})$  with anchor points  $(p_0, t_0) = (x_0, y_0, t_0)$  and  $(p_1, t_1) = (x_1, y_1, t_1)$  consists of the points  $(x, y, t) \in \mathbf{R}^2 \times \mathbf{R}$  that satisfy the inequality constraints  $(x - x_0)^2 + (y - y_0)^2 \leq (t - t_0)^2 \cdot v_{max}^2$  and  $(x - x_1)^2 + (y - y_1)^2 \leq (t_1 - t)^2 \cdot v_{max}^2$ , together with the temporal constraint  $t_0 \leq t \leq t_1$ . Space–time points that satisfy the first of these inequalities are within the future cone since they are closer than the elapsed time times the maximal velocity to the first anchor point, meaning that they can be reached within the available time. Similarly, the space–time points satisfying the second inequality are such that the remaining time is enough to reach the second anchor point. They are therefore in the past cone. In this paper, we assume that a space–time prism is non-empty, which corresponds to the condition  $d(p_0, p_1) \leq (t_1 - t_0) \cdot v_{max}$ , expressing that the second anchor points can be reached from the first within the available time (given the speed limitation).

For a description of many basic properties of prisms, we refer to [9,41]. Here, we only mention that the topological border of the  $C^-(p_0, t_0, v_{max})$  intersects with the topological border of the cone  $C^+(p_1, t_1, v_{max})$  in what we call the rim of the prism  $\mathcal{P}(p_0, t_0, p_1, t_1, v_{max})$ . Figure 1 gives illustrations of one- and two dimensional prisms and their rims. For the one-dimensional case, the rim points  $(x_\ell, t_\ell) = (\frac{(x_1+x_0)-v_{max}(t_1-t_0)}{2}, \frac{v_{max}(t_1+t_0)-(x_1-x_0)}{2v})$  and  $(x_r, t_r) = (\frac{(x_1+x_0)+v_{max}(t_1-t_0)}{2}, \frac{v_{max}(t_1+t_0)-(x_0-x_1)}{2v})$  are easily determined.



**Figure 1.** On the left (a), a one-dimensional prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  with its rim points in green is shown. On the right (b), a space–time prism  $\mathcal{P}(x_0, y_0, t_0, x_1, y_1, t_1, v_{max})$  (in green) is shown for movement in  $\mathbf{R}^2$ . This prism is the intersection of the past cone  $C^+(x_1, y_1, t_1, v_{max})$  (in red) and the future cone  $C^-(x_0, y_0, t_0, v_{max})$  (in blue). The rim  $\rho\mathcal{P}(x_0, y_0, t_0, x_1, y_1, t_1, v_{max})$  is the tilted green ellipse.

#### 4. Random Walk-Based Visit Probability in a One-Dimensional Space–Time Prism

In this section, we describe random walks in a fine-grid network in a one-dimensional space–time prism and the visit probability they induce on the prism. It is convenient to describe such a random walk in a local coordinate system that fits the grid in the prism in which the random walk takes place. In the treatment of the one-dimensional case, we assume that  $x_0 \leq x_1$ . The case  $x_0 \geq x_1$  is completely analogous.

##### 4.1. Random Walks in the Future Cone and Their Local Coordinate System

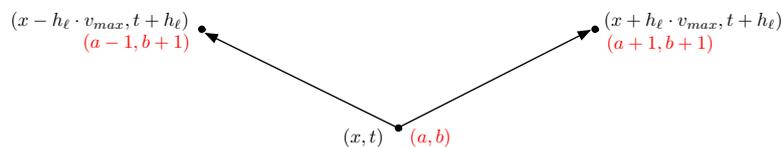
Given a one-dimensional prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$ , we first describe, for any non-zero natural number  $\ell$ , a finite process which generates a random walk that starts at the anchor point  $(x_0, t_0)$  and arrives, after  $\ell$  steps, at the time level  $t_1$  of the second anchor point, while remaining within the bottom

cone  $C^-(x_0, t_0, v_{max})$  of the prism. For a non-zero natural number  $\ell$ , which denotes the number of steps of the random walk, we call  $h_\ell = \frac{t_1 - t_0}{\ell}$  the corresponding step size (or time granularity) of this walk.

The random walk process of  $\ell$  steps starts at level 0 in the anchor point  $(x_0, t_0)$ . The  $i$ -th step of its  $\ell$  steps brings the moving object, performing the random walk, from level  $i - 1$  to level  $i$ . At each level of the process, the moving object starts at some previously reached space–time location  $(x, t)$  and makes, randomly (with equal probability  $\frac{1}{2}$ ), one of the following two moves in the future cone  $C^-(x_0, t_0, v_{max})$  of the prims  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$ :

- (1) from  $(x, t)$  to  $(x - h_\ell \cdot v_{max}, t + h_\ell)$ ; or
- (2) from  $(x, t)$  to  $(x + h_\ell \cdot v_{max}, t + h_\ell)$ .

The two possible steps of the random walk process are illustrated in Figure 2 and they correspond, after the toss of a coin, to moving from  $x$  to the left, respectively to the right, at maximal speed  $v_{max}$  for a time duration  $h_\ell$ . In between levels, we assume these random walk steps to follow linear paths at a maximal speed between these points.



**Figure 2.** The two possible moves in a random walk. The coordinates in the local coordinate system are shown in red, while the Cartesian coordinates are shown in black.

The above-described process arrives, after  $\ell$  steps, at the time level  $t_1$  but it is not guaranteed that it can reach the spatial location  $x_1$  of the second anchor point at that moment. For example, if  $(x_0, t_0) = (0, 0)$ ,  $(x_1, t_1) = (2, 8)$  and  $v_{max} = 1$  (Figure 3 can be used as an illustration), then there is no random walk of  $\ell = 5$  steps from the first to the second anchor point. Indeed, a random walk would take steps of size  $\frac{8}{5}$  in time and in distance. If it would take  $l$  steps to the left and  $r$  steps to the right to reach the second anchor point, we would have  $l + r = \ell = 5$  and  $\frac{8}{5}r - \frac{8}{5}l = 2$ . However, these equations do not have a solution for  $l$  and  $r$  in the natural numbers. Furthermore, the right rim point of this prism has coordinates  $(5, 5)$  and is also not reachable in five steps or fewer from the first anchor point. On the other hand, if we would take  $\ell = 8$  steps, then the second anchor point can be reached by taking 5 steps to the right (arriving in the right rim point), followed by 3 steps to the left, for example.

The ability of a random walk process to reach the second anchor point and, to a lesser extent, the two rim points, is a desirable property, however. The following proposition shows how an appropriate number of levels (and corresponding step size) can be chosen to achieve this. This property only depends on the assumption that  $v_{max}$  and the coordinates of the anchor points are rational numbers. Since the anchor points  $(x_0, t_0)$  and  $(x_1, t_1)$  are the result of measurements (by GPS, for example), this is not unreasonable to assume. For practical purposes,  $v_{max}$  can also be assumed to be a rational number. The proof of this property contains a method to determine a “minimal” number of levels.

**Proposition 1.** For rational  $v_{max}$  and anchor points  $(x_0, t_0)$  and  $(x_1, t_1)$  with rational coordinates, we can determine, from these rational numbers, a “minimal” number of levels  $\ell$  such that a random walk in the bottom cone  $C^-(x_0, t_0, v_{max})$  of the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  can reach the second anchor point after  $\ell$  steps and such that also the two rim points can be visited by random walks of at most  $\ell$  steps.

**Proof.** We assume that  $v_{max}$  and the coordinates of the anchor points  $(x_0, t_0)$  and  $(x_1, t_1)$  are rational numbers. We would like the random walk process, starting in the first anchor point, to be able to reach the second anchor point and the two rim points of the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$ . Let  $\ell$  be a number of

levels (with corresponding step size  $h_\ell$ ) that we are looking for. Let  $d_r$  and  $d_l$  be the number of steps needed to reach the right rim point and the left rim point in a straight motion, respectively. Let  $\bar{d}$  be the “displacement” of the second anchor point, as given by (1), below.

Being able to reach the second anchor point and the two rim points after a finite number of steps in some random walk process corresponds to the existence of natural numbers  $\ell, d_r$  and  $d_l$  and an integer  $\bar{d}$  such that the following six equalities hold:

$$\left\{ \begin{array}{l} x_1 - x_0 = \bar{d} \cdot h_\ell \cdot v_{max} \quad (1) \\ t_1 - t_0 = \ell \cdot h_\ell \quad (2) \\ x_0 - \frac{x_1+x_0}{2} - \frac{v_{max} \cdot (t_1-t_0)}{2} = d_l \cdot h_\ell \cdot v_{max} \quad (3) \\ \frac{(t_1+t_0)}{2} - \frac{x_1-x_0}{2 \cdot v_{max}} - t_0 = d_l \cdot h_\ell \quad (4) \\ \frac{x_1+x_0}{2} + \frac{v_{max} \cdot (t_1-t_0)}{2} - x_0 = d_r \cdot h_\ell \cdot v_{max} \quad (5) \\ \frac{(t_1+t_0)}{2} + \frac{x_1-x_0}{2 \cdot v_{max}} - t_0 = d_r \cdot h_\ell \quad (6) \end{array} \right.$$

Lines (1–2) concern the second anchor point (and (2) is trivially satisfied); lines (3–4) concern the left rim point; and lines (5–6) concern the right rim point. For the left and right rim point, it is important to remark that they can only be reached in one way, namely by moving at maximal speed to the left or the right, respectively, away from  $x_0$ .

From (1), we obtain  $\frac{x_1-x_0}{v_{max} \cdot (t_1-t_0)} = \frac{\bar{d}}{\ell}$ . We remark that we assume that  $x_1 - x_0 \leq v_{max} \cdot (t_1 - t_0)$ , which corresponds to the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  being non-empty. We write the positive rational number  $\frac{x_1-x_0}{v_{max} \cdot (t_1-t_0)}$  as a division free fraction  $\frac{A}{B}$  and take  $\ell = 2 \cdot B$  and  $\bar{d} = 2 \cdot A$ . Therefore, we have  $A \leq B$  or  $\bar{d} \leq \ell$ .

From (3–4), it easily follows that we can take  $d_l = B - A$ , which is a natural number. From (5–6), it easily follows that we can take  $d_r = B + A$ , which is also a natural number.

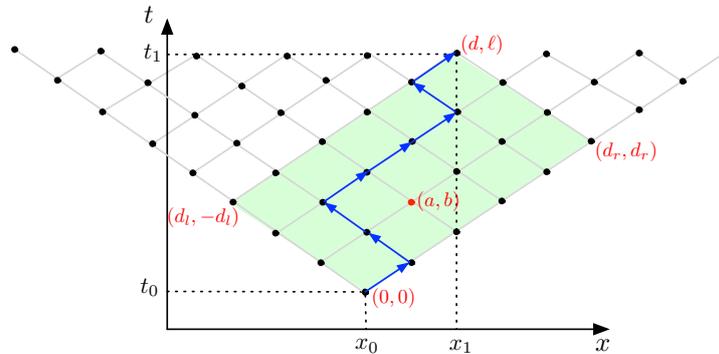
So, it follows that by writing  $\frac{x_1-x_0}{v_{max} \cdot (t_1-t_0)}$  as a division free fraction  $\frac{A}{B}$  we can find the minimal number of levels  $\ell = 2 \cdot B$  that a random walk can take to fit the prism boundaries.  $\square$

The proof of the above proposition gives a “minimum” number of levels  $\ell$  and an associated “displacement”  $\bar{d} = \frac{x_1-x_0}{h_\ell \cdot v_{max}}$  of the second anchor point (corresponding to a shift in the spatial dimension). Obviously, any multiple of the values  $\ell$  and  $\bar{d}$  can also be used for a suitable random walk in the prism. From now on we will work with a number of levels  $\ell$  and a displacement  $\bar{d}$  that are multiples of  $\ell$  and  $\bar{d}$ , that is,  $\ell = n \cdot \ell$  and  $\bar{d} = n \cdot \bar{d}$  for some  $n \in \mathbf{N}_0$ . Thus, Proposition 1 gives rise to a family (parameterized by  $\ell$ ) of fine-grid networks (in the terminology of Burns [30]), on which the the random walk can take place.

Based on the above observations, we can define a local (discrete) grid network and corresponding coordinate system on the bottom cone  $C^-(x_0, t_0, v_{max})$  (and thus on the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$ ), which determine the space–time points in which steps in a random walk from  $(x_0, t_0)$  to  $(x_1, t_1)$  start or end. The local coordinate system assigns the local coordinates  $(0, 0)$  to the first anchor point  $(x_0, t_0)$  and the numbering of further space–time points in a random walk follows the rules: if a grid point with Cartesian coordinates  $(x, t)$  has local coordinates  $(a, b)$ , then  $(x - h_\ell \cdot v_{max}, t + h_\ell)$  and  $(x + h_\ell \cdot v_{max}, t + h_\ell)$  have local coordinates  $(a - 1, b + 1)$  and  $(a + 1, b + 1)$ , respectively. Local coordinates are illustrated in Figure 2, in red. If the second anchor point  $(x_1, t_1)$  has local coordinates  $(d, \ell)$ , we speak about a  $(d, \ell)$ -grid (or  $\ell$ -grid, since  $d$  can be derived from  $\ell$ ) on the prism, when we limit the grid in the future cone to the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$ .

This means that, given some local coordinates  $(a, b)$  in a  $(d, \ell)$ -grid, the corresponding Cartesian coordinates are  $(x_0 + a \cdot h_\ell \cdot v_{max}, t_0 + b \cdot h_\ell)$ . We refer to the first coordinate of the local coordinate system  $(a, b)$  as the displacement  $a$  and to the second coordinate as the level  $b$ . It is important to remark that in a  $(d, \ell)$ -grid the local coordinates  $(a, b)$  are always such that  $a + b$  is even. The value  $a$  varies from  $-b$  to  $b$  in steps of 2 at level  $b$  in the bottom cone of the prism. If we call the value  $r_b(a) = \frac{a+b}{2}$  the rank of  $a$  (at level  $b$ ), for grid points  $(a, b)$ , we can equivalently say that the rank varies from 0 to  $b$  in steps of 1 at level  $b$ .

The grid and local coordinate system are illustrated in Figure 3 with an example in which the number of levels  $\ell = 8$ ,  $d = 2$ ,  $d_l = 3$  and  $d_r = 5$ . The anchor points  $(x_0, t_0)$  and  $(x_1, t_1)$  have local coordinates  $(0, 0)$  and  $(d, \ell) = (2, 8)$ , respectively. The rim points  $(x_\ell, t_\ell)$  and  $(x_r, t_r)$  have local coordinates  $(d_l, -d_l) = (-\frac{\ell-d}{2}, \frac{\ell-d}{2}) = (-3, 3)$  and  $(d_r, d_r) = (\frac{\ell+d}{2}, \frac{\ell+d}{2}) = (5, 5)$ , respectively. Figure 3 also shows an example of a random walk path from the first to the second anchor point where the coin tosses produced the path given by the sequence right-left-left-right-right-right-left-right.



**Figure 3.** The random walk grid network on the bottom cone with the one-dimensional prism in green. Random walks move from level to level between the black points following the grey line segments. The coordinates in red are in the local coordinate system. The blue arrows give an example random walk path from the first to the second anchor point.

We conclude this section by remarking that a random walk from the first to the second anchor point of the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  can be finitely represented, in the local coordinate system, as a sequence  $\langle (0, 0), (a_1, 1), (a_2, 2), \dots, (d, \ell) \rangle$  of grid points, with  $a_{i+1} = a_i \pm 1$ . When a random walk, after  $\ell$  steps, arrives at the second anchor point  $(x_1, t_1)$ , we can then observe that the space–time path resulting from such a random walk (after linear interpolation between the consecutive visited grid points) respects the maximal velocity constraint  $v_{max}$  and is therefore a trajectory within the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  that connects the anchor points. For a precise definition of trajectories in a prism and their velocity, we refer to [51] (with speed bound  $v_{max}$ ).

#### 4.2. Random Walk Visit Probability in Points of a Fine-Grid Network

In this section, the goal is to establish the visit probability in a grid point with local coordinates  $(a, b)$  in a  $(d, \ell)$ -grid on a one-dimensional prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$ . The visit probability of such a grid point is the number of random walks (following the principles, described in Section 4.1) that start at the first anchor point of the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$ , that pass through the grid point  $(a, b)$ , and finish in the second anchor point divided by the total number of random walks from the first to the second anchor point.

We denote the number of random walks that start in the first anchor point and arrive at the grid point with local coordinates  $(a, b)$  by  $\# \uparrow_{paths} (a, b)$ . At level 0, we agree that there is one random walk (of length 0) that arrives at the first anchor point. At level 1, each of the two grid points has one path arriving there. These correspond to the two choices (left or right) that can be made in the first anchor point. This means that we have  $\# \uparrow_{paths} (0, 0) = 1$  and  $\# \uparrow_{paths} (-1, 1) = \# \uparrow_{paths} (1, 1) = 1$ .

For each point, with local coordinates  $(a, b)$ , at a higher level, a random walk arriving there arrives either from  $(a - 1, b - 1)$  or from  $(a + 1, b - 1)$ , provided that the latter points are in the bottom cone. Therefore, we have

$$\# \uparrow_{paths} (a, b) = \# \uparrow_{paths} (a - 1, b - 1) + \# \uparrow_{paths} (a + 1, b - 1),$$

if we consider the terms in this sum that correspond to points outside the bottom cone to be 0. Following the above observations, we recognise that the values  $\# \uparrow^{\text{paths}}(a, b)$  can be read from the well-known binomial triangle (or Pascal’s triangle), since  $\binom{b}{r_b(a)} = \binom{b}{r_{b-1}(a-1)} + \binom{b}{r_{b-1}(a+1)}$  and we obtain

$$\# \uparrow^{\text{paths}}(a, b) = \binom{b}{r_b(a)} = \binom{b}{\frac{a+b}{2}},$$

for  $a, b$  integers with  $0 \leq b, -b \leq a \leq b$  and  $a + b$  even. In terms of generating functions, we can say that  $\# \uparrow^{\text{paths}}(a, b)$  is the coefficient of  $x^{r_b(a)}$  in the expansion of the (rational) polynomial  $(x + \frac{1}{x})^b$  (the proof is straightforward). In this polynomial,  $x$  expresses a move to the right and  $\frac{1}{x}$  corresponds to a move to the left. Figure 4 gives an illustration of the values  $\# \uparrow^{\text{paths}}(a, b)$  for level 0 to level 6 on a random walk grid.

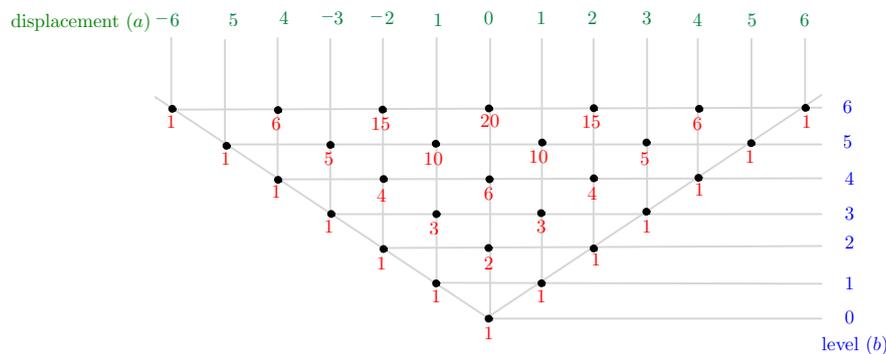


Figure 4. The local grid on the bottom cone with the numbers  $\# \uparrow^{\text{paths}}(a, b)$  in red.

After having established the number  $\# \uparrow^{\text{paths}}(a, b)$  of random walks that arrive at  $(a, b)$ , we need to establish the number of continuations of such walks from  $(a, b)$  to the second anchor point  $(d, \ell)$ . Hereto, it is sufficient to determine the number of “time-reversed” random walks in the future cone that start from the second anchor point and that, moving back in time, arrive in  $(a, b)$ . We denote the number of such downward random walks by  $\# \downarrow^{\text{paths}}(a, b)$ . Each such downward random walk, when time-reversed is an upward random walk from the point  $(a, b)$  to the second anchor point  $(d, \ell)$ .

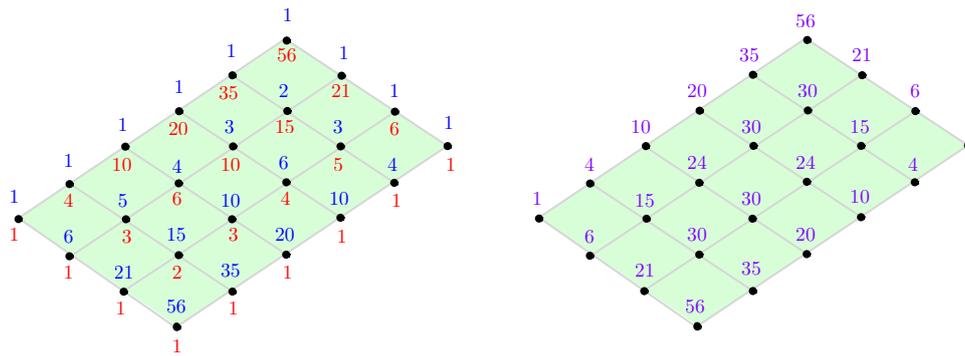
More formally, we can use the mapping  $(x, t) \mapsto (x_1 + x_0 - x, t_1 + t_0 - t)$  (in Cartesian coordinates) or the mapping  $(a, b) \mapsto (d - a, \ell - b)$  (in local coordinates) to map the top cone of the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  onto its bottom cone. When we then apply the earlier method for the bottom cone, we easily verify that  $\# \downarrow^{\text{paths}}(a, b) = \# \uparrow^{\text{paths}}(d - a, \ell - b)$ . The number of paths that come down from the second anchor point and that pass through  $(a, b)$  in the top cone therefore is

$$\# \downarrow^{\text{paths}}(a, b) = \# \uparrow^{\text{paths}}(d - a, \ell - b) = \binom{\ell - b}{r_{\ell - b}(d - a)} = \binom{\ell - b}{\frac{\ell + d}{2} - \frac{a + b}{2}}.$$

Each random walk in the bottom cone from the first anchor point that arrives at  $(a, b)$  can then be combined with such a time-reversed random walk from  $(a, b)$  to the second anchor point to form a random walk going from the first to the second anchor point of the prism, passing through  $(a, b)$ . The total number of random walks in the prism from the first anchor point to the second anchor point, passing through  $(a, b)$ , therefore is the product of  $\# \uparrow^{\text{paths}}(a, b)$  and  $\# \downarrow^{\text{paths}}(a, b)$  and we denote this product by  $\# \updownarrow^{\text{paths}}(a, b)$ . Therefore, we have

$$\# \updownarrow^{\text{paths}}(a, b) = \# \uparrow^{\text{paths}}(a, b) \cdot \# \downarrow^{\text{paths}}(a, b) = \binom{b}{\frac{a+b}{2}} \cdot \binom{\ell - b}{\frac{\ell + d}{2} - \frac{a + b}{2}}$$

The values of  $\# \uparrow^{\text{paths}}(a, b)$ ,  $\# \downarrow^{\text{paths}}(a, b)$  and  $\# \updownarrow^{\text{paths}}(a, b)$  are illustrated in Figure 5.



**Figure 5.** The values  $\# \uparrow^{\text{paths}}(a, b)$  in red,  $\# \downarrow^{\text{paths}}(a, b)$  in blue and  $\# \updownarrow^{\text{paths}}(a, b)$  in purple for the prism of Figure 3.

We remark that the total number of random walk paths in the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  can be read at one of the anchor points, meaning that it is  $\# \updownarrow^{\text{paths}}(0, 0) = \# \updownarrow^{\text{paths}}(d, \ell)$ . In the example of Figure 5, the total number of random walks between the anchor points is  $\# \updownarrow^{\text{paths}}(0, 0) = \# \updownarrow^{\text{paths}}(2, 8) = 56$ .

Now, we are ready to give a definition of “visit probability”, relative to a chosen number of grid levels  $\ell$ . This definition says that the visit probability of a grid point is the fraction of random walk paths in the prism passing through it.

**Definition 2.** Let  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  be a prism and let  $\ell$  be a suitable number of levels for this prism. Let  $(a, b)$  be a point in the  $\ell$ -grid on the prism, given in local coordinates. We define the (one-dimensional) visit probability of  $(a, b)$  (relative to the number of levels  $\ell$ ), denoted  $vp_{1[\ell]}(a, b)$ , to be  $\frac{\# \updownarrow^{\text{paths}}(a, b)}{\# \updownarrow^{\text{paths}}(d, \ell)}$ .

From this definition and the above observations, we immediately obtain the following result.

**Theorem 1.** For a grid point  $(a, b)$  of a  $(d, \ell)$ -grid on a prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$ , we have

$$vp_{1[\ell]}(a, b) = \frac{\binom{b}{r_b(a)} \cdot \binom{\ell-b}{\frac{\ell+d}{2}-r_b(a)}}{\binom{\ell}{\frac{\ell+d}{2}}} = \frac{\binom{b}{\frac{a+b}{2}} \cdot \binom{\ell-b}{\frac{\ell+d}{2}-\frac{a+b}{2}}}{\binom{\ell}{\frac{\ell+d}{2}}}. \tag{7}$$

Using basic properties of binomial coefficients, we can rewrite this in another useful form as

$$vp_{1[\ell]}(a, b) = \frac{\binom{\frac{\ell+d}{2}}{\frac{a+b}{2}} \cdot \binom{\frac{\ell-d}{2}}{\frac{b-a}{2}}}{\binom{\ell}{b}}.$$

When we fix a level  $b$  and sum the visit probability over all  $a$  values at level  $b$ , we expect to obtain 1. This reflects the fact that any random walk from anchor to anchor passes the time slice corresponding to level  $b$  with absolute certainty. The following property verifies this fact.

**Proposition 2.** For any level  $b$ , we have  $\sum_{(a,b) \text{ at level } b} vp_{1[\ell]}(a, b) = 1$ .

**Proof.** The rank of  $a$  at level  $b$  varies from 0 to  $b$ . Some of these ranks correspond to points  $(a, b)$  outside the prism, in which case  $vp_{1[\ell]}(a, b) = 0$ . From (7), we obtain

$$\sum_{(a,b) \text{ at level } b} vp_{1[\ell]}(a, b) = \sum_{r=0}^b \frac{\binom{b}{r} \cdot \binom{\ell-b}{\frac{\ell+d}{2}-r}}{\binom{\ell}{\frac{\ell+d}{2}}},$$

where  $r$  represents the rank. This sum equals 1, since  $\sum_{r=0}^b \binom{b}{r} \cdot \binom{\ell-b}{\frac{\ell+d}{2}-r} = \binom{\ell}{\frac{\ell+d}{2}}$  by the well-known Vandermonde's identity, which can be easily derived by looking at the coefficient of  $x^{\frac{\ell+d}{2}}$  in the expansions of both sides in the equality  $(1+x)^b \cdot (1+x)^{\ell-b} = (1+x)^\ell$ , using the binomial theorem.  $\square$

#### 4.3. Random Walk Visit Probability on the Complete Prism

The visit probability, based on the random walk model, as defined in Definition 2, is specified on the grid points of a  $(d, \ell)$ -grid only. In this section, we discuss how it can be defined on the complete prism. We first discuss the type of distribution associated with this visit probability and the effect of increasing the number of levels of the grid.

##### 4.3.1. The Distribution Corresponding to the Visit Probability

Here, we discuss the type of probability distribution that the visit probability, described by Theorem 1, gives rise to. The first fraction in (7) in Theorem 1, shows that, for each level  $b$ , the expression for  $\text{vp}_{1[\ell]}(a, b)$  corresponds to a hypergeometric distribution in the rank  $r_b(a) = \frac{a+b}{2}$  [52,53]. In fact, if we consider the cumulative density function  $\text{CDF}(a_0 | b, \ell)$  for our situation, we get

$$\sum_{r=0}^{a_0} \frac{\binom{b}{r} \cdot \binom{\ell-b}{\frac{\ell+d}{2}-r}}{\binom{\ell}{\frac{\ell+d}{2}}} = \binom{\ell-b}{\frac{\ell+d}{2}} {}_2F_1 \left[ \begin{matrix} -b; -\frac{\ell+d}{2} \\ 1-b+\frac{\ell-d}{2} \end{matrix} \middle| 1 \right] - \left( \frac{\ell-b}{\frac{\ell+d}{2}-a_0-1} \right) {}_3F_2 \left[ \begin{matrix} 1; 1+a_0-b; 1+a_0-\frac{\ell+d}{2} \\ 2+a_0; 2+a_0-b+\frac{\ell-d}{2} \end{matrix} \middle| 1 \right].$$

The functions  ${}_2F_1$  and  ${}_3F_2$  are so-called hypergeometric functions and we may conclude that we have a hypergeometric distribution [54].

Distributions and also hypergeometric distributions are typically characterized by a number of quantities, such as the mean; the variance; the skewness; and the excess kurtosis [52,53].

In our setting the mean is  $\frac{\ell+d}{2} \cdot \frac{b}{\ell} = r_\ell(d) \cdot \frac{b}{\ell}$ . It is easily verified that the point  $(a, b)$  at level  $b$ , that represents the mean, has a rank  $r_b(a)$  that places it on the diagonal of the prism (given by  $b = \frac{\ell}{d} \cdot a$ , in the local  $(a, b)$ -coordinates).

The variance is  $\frac{\ell^2-d^2}{4\ell^2(\ell-1)} \cdot b \cdot (\ell-b)$ . This variance is tending towards zero towards the anchor points of the prism and reaches its maximum at the level  $b = \frac{\ell}{2}$  (remark that according to Proposition 1 a suitable  $\ell$  is always even).

The skewness is  $\frac{-2(\ell-2b)\sqrt{\ell-1}}{(\ell-2)\sqrt{(\ell^2-d^2)b(\ell-b)}}$ . The skewness is negative for  $0 < b < \frac{\ell}{2}$ , which means that the distribution has a tail (or is stretched) to the left. This is explained by the fact that the left side of the prism, for  $x_0 \leq x_1$ , is further away from the diagonal than its right side. It is zero for  $b = \frac{\ell}{2}$ , which means we have a normal distribution at that level. The skewness is positive for  $\frac{\ell}{2} < b < \ell$ , which means that the distribution has a tail to the right. At the levels  $b = 0$  and  $b = \ell$  the skewness is infinite, which indicates the peak at the unique value at these levels. The excess kurtosis describes how flat or peaky the distribution curve is. In our case, it is a long and complicated expression that we omit here. We only mention that it is negative, which indicates a rather "flat" distribution curve.

##### 4.3.2. The Effect of Refining the Fine-Grid Network

One way to move from a visit probability that is only defined on grid points to one that is defined on the whole prism is to increase the number of levels  $\ell$ , thus approximating each point in the prism eventually by a grid point. In this paragraph, we investigate how the expression for the visit probability  $\text{vp}_{1[\ell]}(a, b)$  in Definition 2 depends on the number of levels  $\ell$ . As an increase in the number of levels, we consider doubling the number of levels and we investigate the effect thereof.

We start with an  $(\bar{d}, \bar{\ell})$ -grid, as described in Section 4.1, and then refine this grid, by for instance iteratively doubling the number of levels, resulting in a  $(2^i \cdot \bar{d}, 2^i \cdot \bar{\ell})$ -grid, with  $i$  a non-zero natural number. This way, we obtain finer and finer grids and an arbitrary point  $(x, t)$  in the prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  can be approximated arbitrarily close. A downside of doubling the number of levels is that the visit probability decreases (and ultimately goes to zero), as the following property shows. This is due to the fact that, when doubling the number of levels, the number of possible paths roughly squares. First, we remark that when the number of levels is doubled, a spatio-temporal point with local coordinates  $(a, b)$  in a  $(d, \ell)$ -grid, will have  $(2a, 2b)$  as local coordinates in a  $(2d, 2\ell)$ -grid. In particular, the second anchor point which has local coordinates  $(d, \ell)$  in the original grid will have local coordinates  $(2d, 2\ell)$  after doubling the number of levels.

**Proposition 3.** For  $\ell, a$  and  $b$  as before, we have  $vp_{1[2\ell]}(2a, 2b) \leq vp_{1[\ell]}(a, b)$ .

The proof of this property is a direct consequence of the well-known estimate

$$\binom{2n}{2k} \sim \binom{n}{k}^2 \sqrt{\pi} \sqrt{\frac{(n-k)k}{n}},$$

from which it follows that  $vp_{1[2\ell]}(2a, 2b) \sim vp_{1[\ell]}(a, b)^2$ , and thus  $vp_{1[2\ell]}(2a, 2b) < vp_{1[\ell]}(a, b)$ , provided that  $0 < vp_{1[\ell]}(a, b) < 1$ . For non-extreme visit probabilities, we thus obtain that the sequence

$$vp_{1[\ell]}(a, b) > vp_{1[2\ell]}(2a, 2b) > vp_{1[4\ell]}(4a, 4b) > \dots > vp_{1[2^k \ell]}(2^k a, 2^k b) > \dots \geq 0$$

converges (by the Weierstrass–Bolzano theorem) to 0, expressing that individual points in finer and finer grids have a visit probability that tends to 0.

This last observation is not surprising. In an infinite set of space–time points, as a prism is, it is to be expected that individual elements have a probability 0. The usual approach to deal with this is to use probability density functions and cumulative density functions, which would allow us to determine the visit probability of a small area around an arbitrary point  $(x, t)$  in a prism  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$ , for instance. Such an area is usually an  $\varepsilon$ -box of type  $[x - \varepsilon, x + \varepsilon] \times [t - \varepsilon, t + \varepsilon]$ . To obtain an exact solution for our case, we can turn to Pearson distributions [55,56], which are the continuous analog of hypergeometric distributions. In our case, we observe that, since the excess kurtosis is negative, the Pearson distribution would be of type 2. Due to the complexity of these Pearson distributions, we present a more practically usable approach in Section 4.3.3.

#### 4.3.3. A Definition of Visit Probability on a One-Dimensional Prism

In this section, we give a practically usable definition of visit probability for all points in a one-dimensional prism. For our definition, we assume that a number of levels  $\ell$  in a grid have been chosen. This number of levels  $\ell$  and the corresponding step size  $h_\ell$  may depend on the application at hand. When the application concerns moving objects that are pedestrians we may want to set  $h_\ell$  to one meter, whereas in the context of moving cars, it may be taken to be larger.

When a number of levels  $\ell$  is fixed, Definition 2 gives us values for the visit probability in grid points on the  $\ell$ -grid on the prism. To determine a value for the visit probability in non-grid points, we use linear interpolation, based on barycentric coordinates in which we use the three closest grid points. This approach gives a more continuous result, as opposed to the voxel-based approach of, for example, References [31,32] and of the random-walk part of [24,26]. Hereto, we partition the grid points into triangles by adding horizontal lines to the grid lines that are already shown in Figure 3. This partition is illustrated in the left part of Figure 6, where also the two types of triangles occurring herein are shown (indicated as ① and ②). We develop the definition of the visit probability of a point with Cartesian coordinates  $(x, t)$  in a type ① in full and remark that for type ② triangles the development is completely analogous. A type ① triangle has corner points with local coordinates

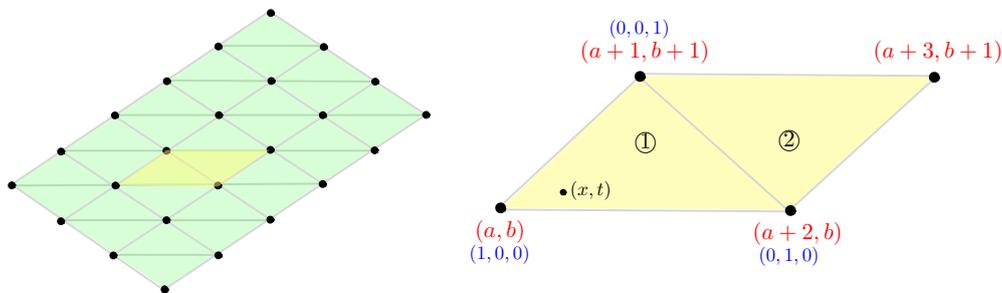
$(a, b)$ ,  $(a + 2, b)$  and  $(a + 1, b + 1)$ . If we give these corner points barycentric coordinates  $(1, 0, 0)$ ,  $(0, 1, 0)$  and  $(0, 0, 1)$ , respectively, then, taking into account that the Cartesian coordinates of a grid point with local coordinates  $(a, b)$  is  $(x_0 + a \cdot h_\ell \cdot v_{max}, t_0 + b \cdot h_\ell)$  (see Section 4.1), we have barycentric coordinates  $(\lambda_1, \lambda_2, \lambda_3)$  for the point  $(x, t)$  if

$$(x, t) = \lambda_1 \cdot (x_0 + a \cdot h_\ell \cdot v_{max}, t_0 + b \cdot h_\ell) + \lambda_2 \cdot (x_0 + (a + 2) \cdot h_\ell \cdot v_{max}, t_0 + b \cdot h_\ell) + \lambda_3 \cdot (x_0 + (a + 1) \cdot h_\ell \cdot v_{max}, t_0 + (b + 1) \cdot h_\ell),$$

with  $0 \leq \lambda_1, \lambda_2, \lambda_3 \leq 1$  and  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ .

We can easily derive that  $\lambda_2 = \frac{1}{2 \cdot h_\ell \cdot v_{max}} \cdot ((x - x_0) + h_\ell \cdot v_{max} \cdot (b - a) - (t - t_0) \cdot v_{max})$ ,  $\lambda_3 = \frac{t - t_0}{h_\ell} - b$  and  $\lambda_1 = 1 - \lambda_2 - \lambda_3 = \frac{1}{2 \cdot h_\ell \cdot v_{max}} \cdot (h_\ell \cdot v_{max} \cdot (b + a + 2) - (x - x_0) + (t - t_0) \cdot v_{max})$ .

Similarly, in a type ② triangle, when we give the grid points with local coordinates  $(a + 1, b + 1)$ ,  $(a + 3, b + 1)$  and  $(a + 2, b)$  barycentric coordinates  $(1, 0, 0)$ ,  $(0, 1, 0)$  and  $(0, 0, 1)$ , respectively, we obtain barycentric coordinates  $(\lambda'_1, \lambda'_2, \lambda'_3)$  for a point  $(x, t)$  with the values  $\lambda'_2 = \frac{1}{2 \cdot h_\ell \cdot v_{max}} \cdot ((x - x_0) - h_\ell \cdot v_{max} \cdot (b + a + 2) + (t - t_0) \cdot v_{max})$ ,  $\lambda'_3 = b + 1 - \frac{t - t_0}{h_\ell}$  and  $\lambda'_1 = 1 - \lambda'_2 - \lambda'_3$ .

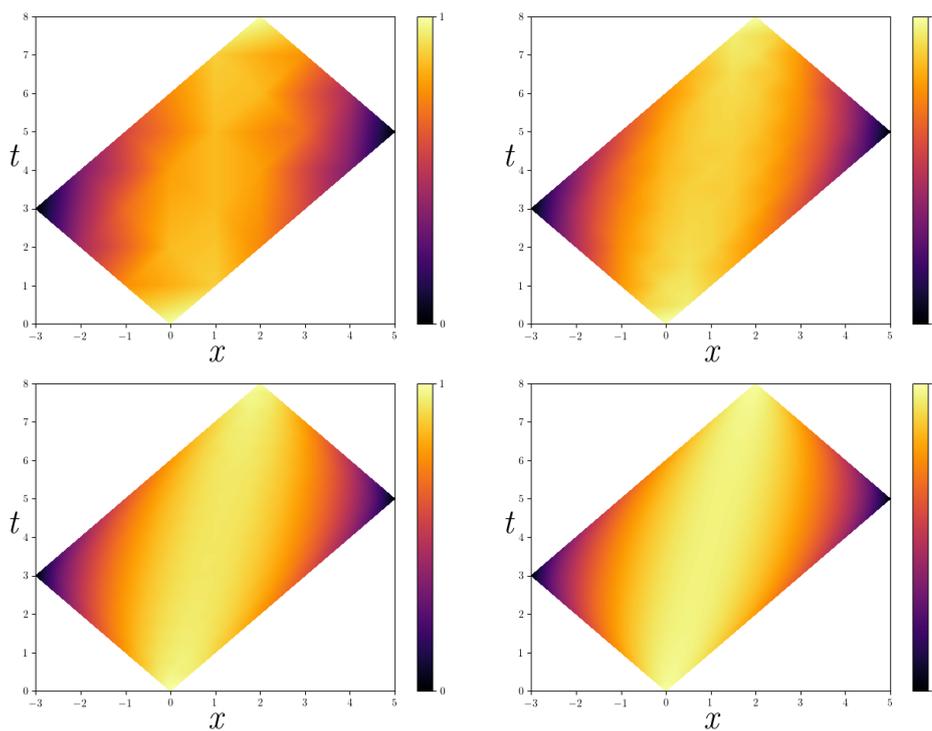


**Figure 6.** The division of a prism in triangles, on the left and the two types of triangles (① and ②) in a prism and their barycentric coordinates, on the right.

Based on these barycentric weights  $\lambda_1, \lambda_2, \lambda_3$  (respectively  $\lambda'_1, \lambda'_2, \lambda'_3$ ) of a point  $(x, t)$  we give the following definition of the visit probability in  $(x, t)$ .

**Definition 3.** Let  $\mathcal{P}(x_0, t_0, x_1, t_1, v_{max})$  be a prism and let  $\ell$  be a suitable number of levels for this prism. If  $(x, y)$  is a point in a type ① triangle, as shown in Figure 6, then we define  $vp_{1[\ell]}(x, y)$  to be  $\lambda_1 \cdot vp_{1[\ell]}(a, b) + \lambda_2 \cdot vp_{1[\ell]}(a + 2, b) + \lambda_3 \cdot vp_{1[\ell]}(a + 1, b + 1)$ . Let  $(x, y)$  be a point in a type ② triangle, as shown in Figure 6. Then we define  $vp_{1[\ell]}(x, y)$  to be  $\lambda'_1 \cdot vp_{1[\ell]}(a + 1, b + 1) + \lambda'_2 \cdot vp_{1[\ell]}(a + 3, b + 1) + \lambda'_3 \cdot vp_{1[\ell]}(a + 2, b)$ . □

Figure 7 shows the barycentric-based visit probability for the prism of Figure 3, where we take the number of levels  $\ell$  to be respectively 8, 16, 32 and 64. These illustrations show that the areas around the anchor points have a high probability of being visited, while the areas around the rim points are assigned a very low probability. In general, the diagonal of the prism shows the highest visit probability.



**Figure 7.** The barycentric-based visit probability for the prism of Figure 3 for  $\ell$  values of 8, 16, 32 and 64.

## 5. Random Walk-Based Visit Probability in a Two-Dimensional Space–Time Prism

In this section, we develop the random-walk based visit probability for prisms that describe movement in a two-dimensional space. The main difference with the one-dimensional case is that the fine-grid network does not fill the complete prism, whereas in the one-dimensional case it does. The fine-grid network takes the form of a pyramidal prism in this case.

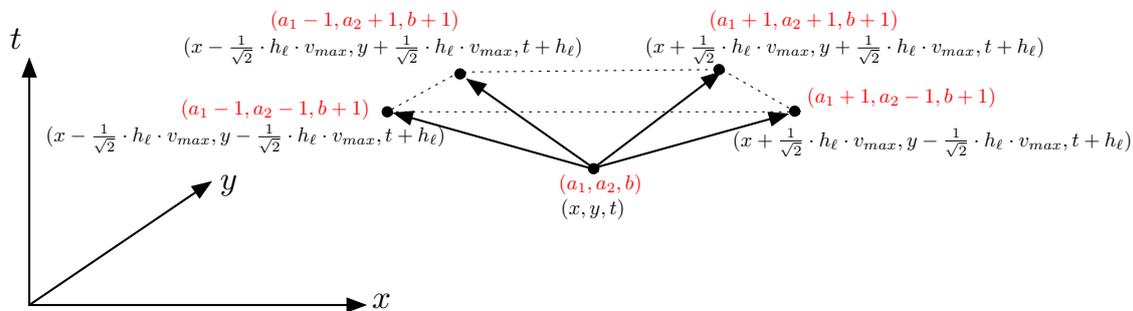
### 5.1. Random Walks in the Future Cone and Their Local Coordinate System

To facilitate the description of the two-dimensional case, we assume that the prism  $\mathcal{P}(x_0, y_0, t_0, x_1, y_1, t_1, v_{max})$  has its first anchor point at the origin and the second anchor point above the (positive) spatial diagonal given by the equations  $y = x$  and  $x \geq 0$ , that is,  $(x_0, y_0, t_0) = (0, 0, 0)$  and  $(x_1, y_1, t_1) = (x_1, x_1, t_1)$ , with  $x_1 \geq 0$ . This assumption is without loss of generality since any prism can be transformed to this particular prism by an isometric transformation of the  $(x, y, t)$ -space (with the appropriate choice of  $t_1$ ).

When we intersect the prism  $\mathcal{P}(0, 0, 0, x_1, x_1, t_1, v_{max})$ , with the plane in the  $(x, y, t)$ -space given by the equation  $y = x$ , we obtain a one-dimensional prism and we can use Proposition 1 to obtain a suitable minimum number of levels  $\ell$  (of which we can use any multiple) to be able to reach the second anchor from the first in  $\ell$  steps.

We start by describing a finite process that generates, in  $\ell$  steps, a random walk within the future cone of the prism. Let  $h_\ell = \frac{t_1 - t_0}{\ell}$  be the corresponding step size of this process. A random walk of  $\ell$  steps starts at the origin when we are at level 0. The  $i$ -th of these  $\ell$  steps brings the moving object from level  $i - 1$  to level  $i$  in the random walk process. At each level of the process, the moving object, that performs the random walk, starts at some previously reached space–time location  $(x, y, t)$  and tosses two independent coins: the first will decide whether to move in the positive or negative  $x$ -direction (North or South) and the second coin does the same for the  $y$ -direction (East or West). The steps in the  $x$  and  $y$ -direction are of spatial size  $\frac{1}{\sqrt{2}} \cdot h_\ell \cdot v_{max}$ , such that when they are combined (added), the resulting step size is  $h_\ell \cdot v_{max}$ . Starting from  $(x, y, t)$ , the following four possible space–time locations are equally likely outcomes of a step: (1)  $(x + \frac{1}{\sqrt{2}} \cdot h_\ell \cdot v_{max}, y + \frac{1}{\sqrt{2}} \cdot h_\ell \cdot v_{max}, t + h_\ell)$ ; (2)  $(x + \frac{1}{\sqrt{2}} \cdot$

$h_\ell \cdot v_{max}, y - \frac{1}{\sqrt{2}} \cdot h_\ell \cdot v_{max}, t + h_\ell$ ); (3)  $(x - \frac{1}{\sqrt{2}} \cdot h_\ell \cdot v_{max}, y + \frac{1}{\sqrt{2}} \cdot h_\ell \cdot v_{max}, t + h_\ell)$ ; and (4)  $(x - \frac{1}{\sqrt{2}} \cdot h_\ell \cdot v_{max}, y - \frac{1}{\sqrt{2}} \cdot h_\ell \cdot v_{max}, t + h_\ell)$ . These four possible moves are in the NE, SE, SW or NW directions (in terms of a compass) and are illustrated in Figure 8. Since we have located the second anchor point above the positive diagonal, initially at least, option (1) is a move towards the second anchor point.



**Figure 8.** The four possible moves in a two-dimensional random walk. The local coordinates are shown in red.

In the two-dimensional case, we want to introduce a fine-grid network, or grid, for short, on which the random walk takes place together with a local coordinate system. When we have chosen, in accordance with Proposition 1, a number of levels  $\ell$ , we give the first anchor point the local  $(a_1, a_2, b)$ -coordinates  $(0, 0, 0)$ . Further points in the random walk grid get assigned local coordinates as shown in Figure 8 in red. Proceeding this way, the second anchor point will have local coordinates  $(d_1, d_2, \ell)$  and we speak of a  $(d_1, d_2, \ell)$ -grid on the prism (or  $\ell$ -grid, for short). It is easily verified that on such a grid with local coordinates  $(a_1, a_2, b)$  we have that  $a_1 + a_2, a_1 + b$  and  $a_2 + b$  are all even.

## 5.2. Random Walk Visit Probability in Grid Points

The following lemma gives an expression for the count of the number of random walk paths that start from the first anchor point and arrive in a point in the bottom cone of the prism with local coordinates  $(a_1, a_2, b)$ . We call this number  $\# \uparrow^{\text{paths}}(a_1, a_2, b)$ , in analogy with the one-dimensional case.

**Lemma 1.** For a grid point  $(a_1, a_2, b)$  of a  $(d_1, d_2, \ell)$ -grid on a prism  $\mathcal{P}(x_0, y_0, t_0, x_1, y_1, t_1, v_{max})$ , we have  $\# \uparrow^{\text{paths}}(a_1, a_2, b) = \binom{b}{a_1+b} \cdot \binom{b}{a_2+b}$ .

**Proof.** A random walk that reaches  $(a_1, a_2, b)$  comes from  $(a_1 - 1, a_2 - 1, b - 1)$ ,  $(a_1 - 1, a_2 + 1, b - 1)$ ,  $(a_1 + 1, a_2 - 1, b - 1)$  or  $(a_1 + 1, a_2 + 1, b - 1)$ . We therefore have  $\# \uparrow^{\text{paths}}(a_1, a_2, b) = \# \uparrow^{\text{paths}}(a_1 - 1, a_2 - 1, b - 1) + \# \uparrow^{\text{paths}}(a_1 - 1, a_2 + 1, b - 1) + \# \uparrow^{\text{paths}}(a_1 + 1, a_2 - 1, b - 1) + \# \uparrow^{\text{paths}}(a_1 + 1, a_2 + 1, b - 1)$ , if we consider the terms in this sum that correspond to points outside the prism to be 0. It is clear that for each level  $b$ ,  $\# \uparrow^{\text{paths}}(b, b, b) = \# \uparrow^{\text{paths}}(b, -b, b) = \# \uparrow^{\text{paths}}(-b, b, b) = \# \uparrow^{\text{paths}}(-b, -b, b) = 1$ , since there are unique paths of maximal speed to these extremal points.

We claim that  $\# \uparrow^{\text{paths}}(a_1, a_2, b)$  is the coefficient of  $x^{a_1} \cdot y^{a_2}$  in the expansion of the polynomial  $(x + \frac{1}{x})^b \cdot (y + \frac{1}{y})^b$ . We prove this by induction on  $b$ . For the level  $b = 0$ , we have one path of length 0 and thus  $\# \uparrow^{\text{paths}}(0, 0, 0) = 1$ . Clearly, also the coefficient of  $x^0 \cdot y^0$  in the expansion  $(x + \frac{1}{x})^0 \cdot (y + \frac{1}{y})^0$  is 1. We assume that the claim is true for level  $b - 1$  and prove it for level  $b$ . Since  $(x + \frac{1}{x})^b \cdot (y + \frac{1}{y})^b = (x + \frac{1}{x})^{b-1} \cdot (y + \frac{1}{y})^{b-1} \cdot (x + \frac{1}{x}) \cdot (y + \frac{1}{y})$ , we obtain from the induction hypothesis that

$$(x + \frac{1}{x})^b \cdot (y + \frac{1}{y})^b = (\sum \# \uparrow^{\text{paths}}(a_1, a_2, b - 1) x^{a_1} \cdot y^{a_2}) \cdot (x + \frac{1}{x}) \cdot (y + \frac{1}{y}),$$

where the sum ranges over all  $a_1, a_2$  with  $-b + 1 \leq a_1, a_2 \leq b - 1$ , and  $a_1 + a_2, a_1 + b - 1$  and  $a_2 + b - 1$  all even. When we look for the coefficient of  $x^{a_1} \cdot y^{a_2}$  in  $(x + \frac{1}{x})^b \cdot (y + \frac{1}{y})^b$  and we consider that

$x^{a_1} \cdot y^{a_2} = x \cdot x^{a_1-1} \cdot y^{a_2} = \frac{1}{x} \cdot x^{a_1+1} \cdot y^{a_2} = y \cdot x^{a_1} \cdot y^{a_2-1} = \frac{1}{y} \cdot x^{a_1} \cdot y^{a_2+1}$ , we obtain that the coefficient of  $x^{a_1} \cdot y^{a_2}$  is  $\# \uparrow^{\text{paths}}(a_1 - 1, a_2 - 1, b - 1) + \# \uparrow^{\text{paths}}(a_1 - 1, a_2 + 1, b - 1) + \# \uparrow^{\text{paths}}(a_1 + 1, a_2 - 1, b - 1) + \# \uparrow^{\text{paths}}(a_1 + 1, a_2 + 1, b - 1)$ , which is  $\# \uparrow^{\text{paths}}(a_1, a_2, b)$ , as we have observed before. This completes the proof of the claim.

Next, we see that  $(x + \frac{1}{x})^b \cdot (y + \frac{1}{y})^b = (\sum_{i=0}^b \binom{b}{i} x^i \cdot (\frac{1}{x})^{b-i}) \cdot (\sum_{j=0}^b \binom{b}{j} y^j \cdot (\frac{1}{y})^{b-j}) = (\sum_{i=0}^b \binom{b}{i} x^{2i-b}) \cdot (\sum_{j=0}^b \binom{b}{j} y^{2j-b})$ . The coefficient of  $x^{a_1} \cdot y^{a_2}$  therefore is  $(\frac{b}{\frac{a_1+b}{2}}) \cdot (\frac{b}{\frac{a_2+b}{2}})$  and we obtain  $\# \uparrow^{\text{paths}}(a_1, a_2, b) = (\frac{b}{\frac{a_1+b}{2}}) \cdot (\frac{b}{\frac{a_2+b}{2}})$ . This completes the proof.  $\square$

The above lemma shows that  $\# \uparrow^{\text{paths}}(a_1, a_2, b) = \binom{b}{r_b(a_1)} \cdot \binom{b}{r_b(a_2)} = \# \uparrow^{\text{paths}}(a_1, b) \cdot \# \uparrow^{\text{paths}}(a_2, b)$ , thus linking the one- and two-dimensional cases.

Similar to the one-dimensional case, we use the notation  $\# \downarrow^{\text{paths}}(a_1, a_2, b)$  to count the number of (time-reversed) random walk paths from the second anchor point downward to the grid point with local coordinates  $(a_1, a_2, b)$ . Using the mapping  $(a_1, a_2, b) \mapsto (d_1 - a_1, d_2 - a_2, \ell - b)$ , we can see that  $\# \downarrow^{\text{paths}}(a_1, a_2, b) = \# \uparrow^{\text{paths}}(d_1 - a_1, d_2 - a_2, \ell - b)$ . Finally, we denote by  $\# \updownarrow^{\text{paths}}(a_1, a_2, b)$  the total number of random walk paths from the first to the second anchor, passing through  $(a_1, a_2, b)$  and this number is  $\# \uparrow^{\text{paths}}(a_1, a_2, b) \cdot \# \downarrow^{\text{paths}}(a_1, a_2, b)$ . Again, like in the one-dimensional case,  $\# \updownarrow^{\text{paths}}(0, 0, 0) = \# \updownarrow^{\text{paths}}(d_1, d_2, \ell)$  is the total number of random walk paths in the prism from the first to the second anchor.

In analogy with Definition 2, we define the visit probability for grid points in the two-dimensional case.

**Definition 4.** Let  $\mathcal{P}(x_0, y_0, t_0, x_1, y_1, t_1, v_{max})$  be a prism and let  $\ell$  be a suitable number of levels for this prism. Let  $(a_1, a_2, b)$  be a grid point given in the local coordinate system. We define the (two-dimensional) visit probability of  $(a_1, a_2, b)$  (relative to the number of levels  $\ell$ ), denoted  $vp_{2[\ell]}(a_1, a_2, b)$ , to be  $\frac{\# \updownarrow^{\text{paths}}(a_1, a_2, b)}{\# \updownarrow^{\text{paths}}(d_1, d_2, \ell)}$ .  $\square$

From Lemma 1, we immediately obtain the following an expression for the two-dimensional random walk-based visit probability.

**Theorem 2.** For a grid point  $(a_1, a_2, b)$  of a  $(d_1, d_2, \ell)$ -grid on a prism  $\mathcal{P}(x_0, y_0, t_0, x_1, y_1, t_1, v_{max})$ , we have

$$vp_{2[\ell]}(a_1, a_2, b) = vp_{1[\ell]}(a_1, b) \cdot vp_{1[\ell]}(a_2, b).$$

**Proof.** Using Definition 4 and Lemma 1, we therefore obtain

$$vp_{2[\ell]}(a_1, a_2, b) = \frac{(\frac{b}{\frac{a_1+b}{2}}) \cdot (\frac{b}{\frac{a_2+b}{2}}) \cdot (\frac{\ell-b}{\frac{d_1+\ell}{2} - \frac{a_1+b}{2}}) \cdot (\frac{\ell-b}{\frac{d_2+\ell}{2} - \frac{a_2+b}{2}})}{(\frac{\ell}{\frac{d_1+\ell}{2}}) \cdot (\frac{\ell}{\frac{d_2+\ell}{2}})},$$

which equals

$$\frac{(\frac{b}{\frac{a_1+b}{2}}) \cdot (\frac{\ell-b}{\frac{d_1+\ell}{2} - \frac{a_1+b}{2}})}{(\frac{\ell}{\frac{d_1+\ell}{2}})} \cdot \frac{(\frac{b}{\frac{a_2+b}{2}}) \cdot (\frac{\ell-b}{\frac{d_2+\ell}{2} - \frac{a_2+b}{2}})}{(\frac{\ell}{\frac{d_2+\ell}{2}})}$$

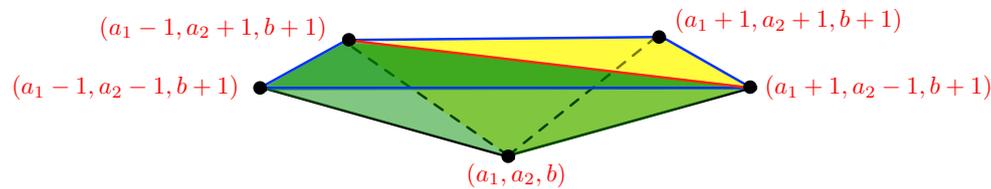
or  $vp_{1[\ell]}(a_1, b) \cdot vp_{1[\ell]}(a_2, b)$ . This completes the proof.  $\square$

This theorem exposes a link between the one- and the two-dimensional cases.

### 5.3. Random Walk Visit Probability on the Complete Prism

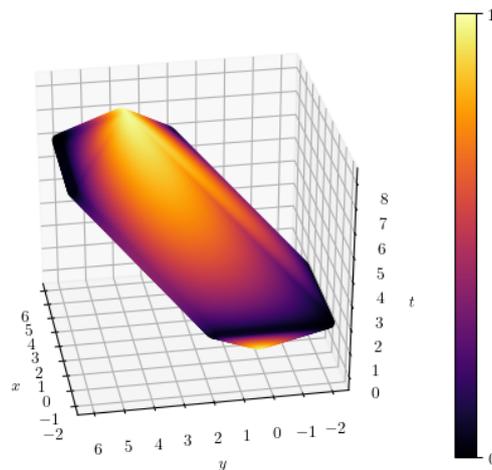
As for the one-dimensional case, we can extend the visit probability to the complete prism. However, we should note that the fine-grid network does not cover the complete prism, but rather a pyramidal sub prism of it. Without going in all the technical details for this case, we illustrate

how the visit probability works in the two-dimensional setting. The pyramids between two layers, as illustrated in Figure 8, are split into two tetrahedra. Hereto, we split the ground plane of the pyramid into two triangles, as illustrated by the red line between the green and yellow triangles in Figure 9. We then use barycentric interpolation within these tetrahedra to obtain a continuous version of the visit probabilities on the fine-grid network.



**Figure 9.** The pyramid is split into two tetrahedra (along the red line in the ground plane of the pyramid).

Figure 10 gives an illustration of a visit probability heat map on a pyramidal shaped prism. The prism anchor points are  $(0, 0, 0)$  and  $(4, 4, 8)$ ,  $v_{max} = 1$  and  $\ell = 256$  is used in this example. Again, we see that the vicinity of the anchor points gives a high visit probability, whereas the border of the rim plane has the lowest visit probability. According to the hypergeometric distribution, the diagonal contains the most likely points within the prism. Obviously, the complete classical two-dimensional prism is larger. As in [26], we can again use some interpolation method to fill the gap between the pyramidal and the classical prism.



**Figure 10.** The barycentric-based visit probability for the pyramidal shaped prism with anchor points  $(0, 0, 0)$  and  $(4, 4, 8)$ , for  $v_{max} = 1$  and the number of levels  $\ell$  being 256.

## 6. Examples of Applications of the Visit Probability Measure

In this section, we give some example scenarios of how space–time prisms enriched with a visit probability measure could be used in practical applications. Our application domains are about human movement in on a road network and animal movement in two-dimensional space. The problems involve questions like: “With which level of confidence has a moving object visited some location?” and “Which is the more likely path followed by a moving object?”.

### 6.1. Human Movement and Interaction

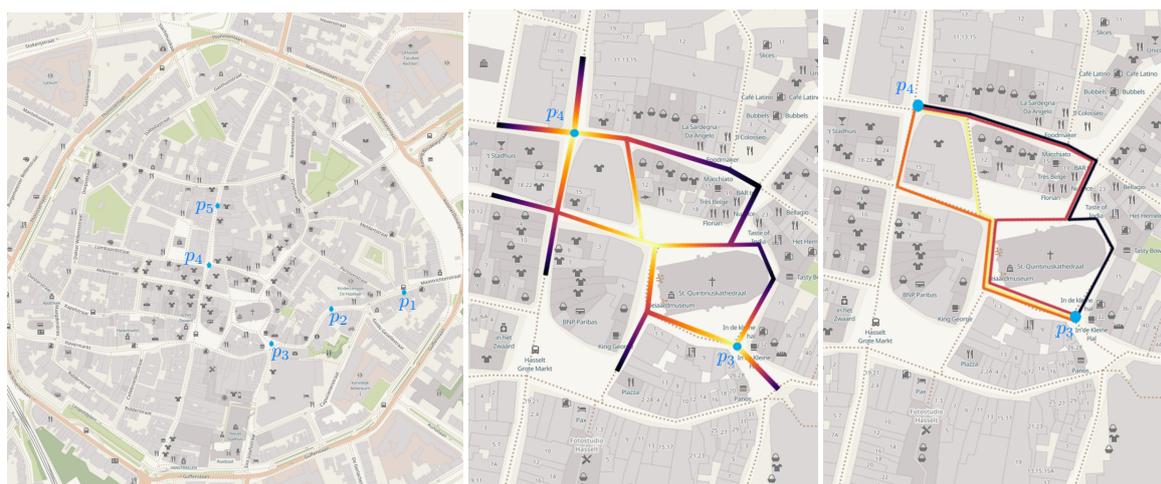
Human activities are encompassed within spatio-temporal constraints, such as (physical or law-imposed) speed limitations and the presence of obstacles [47]. The space–time prism model has

already successfully been applied to model such situations. By overlaying a probabilistic framework on the space–time prisms, this model can be enriched [8,24,26].

For example, if we consider the movement of people (like pedestrians) on a road network in a city, we first remark that on a road network, we basically are in the one-dimensional case, that was discussed extensively in this paper. The one-dimensional prisms can be “folded” over the road network, following its possible ramifications, as is explained in depth in [8]. We assume that the location of these people is registered (for instance, by GPS) at some regular moments in time. Then, we can use the binomial random-walk based visit probability as we demonstrate in the following example of a pedestrian moving in a city.

### 6.1.1. Movement Modelling

For this example, we collected data during a (pedestrian) walk in the city center of Hasselt, Belgium. Five measurements, collected by Google Fitness tracker, were taken at regular time intervals and they are shown in the first part of Figure 11 as  $p_1$ ,  $p_2$ ,  $p_3$ ,  $p_4$  and  $p_5$ . In this map of Hasselt, the paths that a pedestrian can follow are shown as brown dotted lines. The distance covered by our measured walk was 550 m, and the maximal walking speed of the pedestrian was calculated at 1.185 m/s (or 4.2 km/h), based on previously recorded walks. The walk was timed at 8.21 min with measurements taken at every two minutes interval (starting at the 0th minute).



**Figure 11.** An illustration of the “visit probability” between two anchor points for a pedestrian on a street network in the city of Hasselt. The left most map shows the measured data points during the walk. The map in the middle focuses on two measured points and the spatial projection of the one-dimensional prism between them, colored with visit probabilities. The rightmost figure shows the paths that could be taken and how probable they are in a color gradient, with bright yellow being the most probable and dark blue the least probable.

We focus on the data points  $p_3$  and  $p_4$  measured at the 4th and the 6th minute. The points have longitude–latitude coordinates (5.33901, 50.92954) and (5.33768, 50.93067), respectively, given in decimal degrees. We will consider them as the first and second anchor point of the prism to be constructed.

The second part of Figure 11 shows the spatial projection (or PPA) of the one-dimensional prism (“folded” over the road network, following its possible ramifications, using the algorithm of [8]), augmented with colors that reflect the visit probability of these spatial points. The colour scale used is the same as in Figures 7 and 10 and the visit probability is computed as the average visit probability of all space–time points above the spatial location. This middle figure shows us that certain deviations into side roads are unlikely and that, for example, visiting the north side of the cathedral (the building with the cross on it) is highly likely.

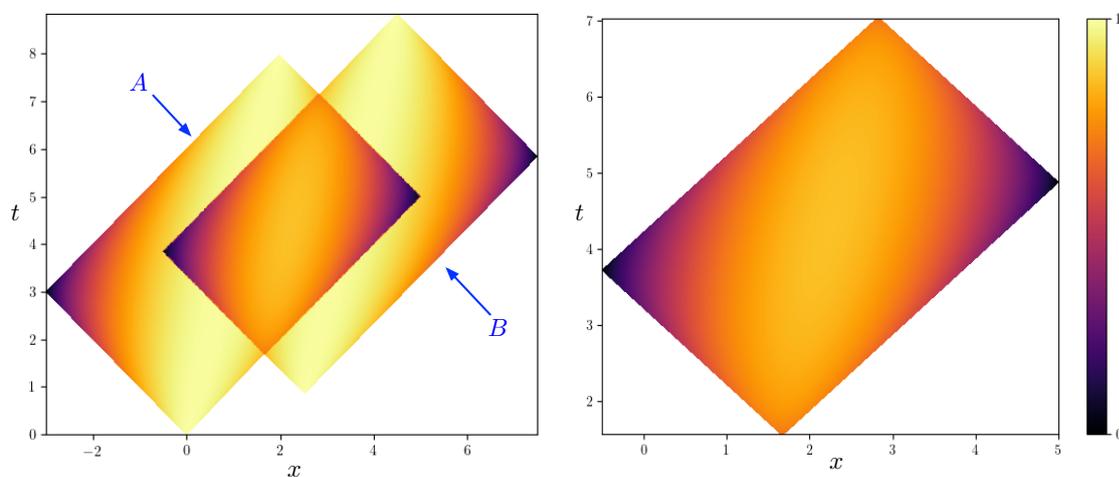
In the rightmost part of Figure 11, we consider the possible direct paths between  $p_3$  and  $p_4$  (without excursions into side roads) and we assign the average visit probability (taken over all their spatial points) as colors to them. This figure shows that the dark blue path (on the right side) is the least likely, whereas the yellow path that almost diagonally connects  $p_3$  and  $p_4$  is there more likely. The yellow path has an average probability of 0.8418, the orange path meanwhile has an average probability of 0.63, the red path has an average probability of 0.4807, while the final possible path, colored dark blue has a very low average probability of 0.1699. These probabilities could be used to estimate the likelihood of this pedestrian having been at some place or event.

We remark that these data were collected by one of the authors while on a walk. The map as well as the measured data points were recorded with the help of OpenStreetMap [57]. The maximal walking speed of the pedestrian was calculated from the compiled data collected by the authors' fitness tracking application over the course of a year.

### 6.1.2. Human Interaction

Let us now consider the question of when and where two moving objects  $A$  and  $B$  could have met (while also giving a confidence level). Assuming that  $A$  and  $B$  move independently, we can use the product of their (binomial random-walk based) visit probability as a measure of "meeting probability".

We illustrate this situation in Figure 12, where only one prism for each of the moving object  $A$  and  $B$  is shown (usually from a longer chain of prisms). Here, the anchor points of  $A$  are  $(x_0, y_0) = (0, 0)$  and  $(x_1, y_1) = (1.75, 8)$  and those of  $B$  are  $(x'_0, y'_0) = (2.5, 1)$  and  $(x'_1, y'_1) = (4.25, 9)$  and  $v_{max}$  is assumed to be 1. Figure 12 shows on the left, outside the intersection of the prisms of  $A$  and  $B$ , the binomial random-walk probability (based on 256 levels) that we discussed in this paper. However, inside the intersection of the prisms of  $A$  and  $B$ , the meeting probability (defined as the product of their visit probabilities) is shown. In this example it varies between 0.0004 and 0.6427.



**Figure 12.** An illustration of the "meeting probability" of two moving objects that move on a straight line (or a road network). The right part focusses on the intersection of the two prisms.

This information can be used to determine more likely places and moments in time where  $A$  and  $B$  could have met. Obviously, this problem is related to the alibi query [20], which asks whether or not two prisms are disjoint, but clearly, in the presence of visit probability richer information can be obtained.

The previous query focusses on two moving objects, but when, for instance, historical data about many moving objects is present, we could ask questions about the optimal location of new facilities. For example, if we plan to open a new restaurant in a shopping area of a city, we might be interested in locations where the cumulative visit probability is higher or highest at certain moments of the day.

## 6.2. Animal Movement and Interaction

The study of the movement and the interaction of animals plays a significant role in improving our understanding of different species and their group dynamics, which in turn is of great importance in the preservation and conservation of their ecosystems (see, for example, References [25,32,58] and references in the Introduction). If we consider the scenario of tracking the movement of animals with sensors (such as Argos). Typically, to conserve battery life in the sensors, the signals are sent in intervals of hours or days, whereas even a 15 min frequency is used nowadays. In order to provide an analysis (real-time or not) of animal behavior and interaction the unsampled time intervals need to be accounted for. Ideally, this should be accompanied by statistical analysis, for example, in terms of visit probability.

For this case study, we will consider the movement of an African Buffalo *Syncerus caffer*. The buffaloes in the study have been tagged with GPS tracking devices with a signal being sent by the tag at regular intervals, in this case every 60 min. We consider one day (17 February 2005) in the life of the buffalo “Queen” with the tag T8. The GPS tracking converted to the spatial coordinates can be seen in the left part of Figure 13. The african buffaloes have been studied extensively and have been determined to have a maximal speed of 13.33–15.27 m/s and an average walking speed of 1.42 m/s [59]. Since this is a simple case study to demonstrate the binomial random walk model proposed in this paper, we will consider the maximal speed of the buffalo to be 15 m/s. Given the velocity bound, we can construct the two-dimensional prisms between the measured data points as seen in the right part of Figure 13. We can see from the image all possible areas the buffalo could have been between measured data points, given its maximal velocity. Since this example has purely illustrative purposes, we overestimated the maximal velocity in order to show the PPAs more clearly. This example serves to illustrate, in the right part of Figure 13, the likelihood of places this buffalo has visited. This might have an application in studying the probability of this animal having visited some food resources. The same model can be applied in other situations in animal ecology, such as the determination of the risk of an animal has been in contact with other diseased animals (for instance, an animal colony infected by influenza).



**Figure 13.** The above figure illustrates the application of “visit probability” in the study of animal movement. The figure to the left indicates the movement of a GPS tagged African Buffalo in a day. The figure to the right shows the visit probabilistic model applied to compensate for the missing data points.

The data used in the above use case was of 17 February 2005 and was published in *Movebank* [60] by Paul Cross.

## 7. Conclusions

We have studied a framework to augment classical space–time prisms with an internal structure in terms of a visit probability function. This function maps space–time points within the prism to a value in the unit interval. This value expresses their likelihood of being visited by a path connecting

the anchor points of the prism. The proposed mathematical framework is based on a binomial random walk within one- and two-dimensional space–time prisms. Our random walk takes place on a fine-grid network within the prism and in studying it, we make no assumptions on the random walks, we impose no distributions (as is done in [26]), we have no truncations and we do not introduce a bias towards the second anchor point. By simply defining the random walk on a grid network, we arrive at the conclusion that random walk-based visit probability in space–time prisms corresponds to a hypergeometric distribution.

This work can be expanded and deepened in a number of directions. Firstly, our random walk has the choice between two directions (going left or right). This could be extended, as already suggested in [24], to include standing still. With these three options, a similar approach can be followed, which would result in a trinomial random-walk based visit probability, which is yet to be investigated further. Secondly, both in the one- and two-dimensional cases, there remains the problem of moving from a discrete probability distribution (on a fine-grid network) to a probability distribution on the complete prism. Pearson distributions [55,56] need to be further investigated in this context. Thirdly, in the two-dimensional case, the fine-grid network corresponds to a pyramidal prism which is part of the classical prism. It remains to be investigated how a random walk on the complete classical prism can be defined. In our current proposal the random walk is based on two orthogonal directions. One way to cover the complete classical prisms could be to increase the number of possible directions. Finally, the proposed visit probability should be incorporated in (formal) query systems, such as the one proposed in [9], to allow for queries that deal with probability-based space–time objects. The same holds for trajectory analysis systems.

**Author Contributions:** Both authors contributed equally to all aspects of this paper. Figures 7, 10 and 12 were produced by Deepak Elias. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors would like to thank Bart Cleuren for useful discussions on the topic of the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Neutens, T.; Van de Weghe, N.; Witlox, F.; De Maeyer, P. A three-dimensional network-based space–Time prism. *J. Geogr. Syst.* **2008**, *10*, 89–107. [\[CrossRef\]](#)
2. Neutens, T.; Delafontaine, M.; Schwanen, T.; van de Weghe, N. The relationship between opening hours and accessibility of public service delivery. *J. Trans. Geogr.* **2012**, *25*, 128–140. [\[CrossRef\]](#)
3. Neutens, T.; Delafontaine, M.; Scott, D.M.; De Maeyer, P. An analysis of day-to-day variations in individual space-time accessibility. *J. Trans. Geogr.* **2012**, *23*, 81–91. [\[CrossRef\]](#)
4. Kwan, M.P.; Hong, X.D. Network-based constraints-oriented choice set formation using GIS. *J. Geogr. Syst.* **1998**, *5*, 139–162.
5. Kwan, M.P. Interactive geovisualization of activity–travel patterns using three-dimensional geographical information systems: A methodological exploration with a large data set. *Transp. Res. Part C Emerg. Technol.* **2000**, *8*, 185–230. [\[CrossRef\]](#)
6. Yu, H.; Shaw, S.L. Exploring potential human activities in physical and virtual spaces: a spatio-temporal GIS approach. *Int. J. Geogr. Inf. Sci.* **2008**, *22*, 409–430. [\[CrossRef\]](#)
7. Ratcliffe, J.H. A Temporal Constraint Theory to Explain Opportunity-Based Spatial Offending Patterns. *J. Res. Crime Delinq.* **2006**, *43*, 261–291. [\[CrossRef\]](#)
8. Kuijpers, B.; Othman, W. Modeling uncertainty of moving objects on road networks via space-time prisms. *Int. J. Geogr. Inf. Sci.* **2009**, *23*, 1095–1117. [\[CrossRef\]](#)
9. Kuijpers, B.; Othman, W. Trajectory databases: Data models, uncertainty and complete query languages. *J. Comput. Syst. Sci.* **2010**, *76*, 538–560. [\[CrossRef\]](#)
10. Neutens, T.; Schwanen, T.; Witlox, F.; Maeyer, P.D. My space or your space? Towards a measure of joint accessibility. *Comput. Environ. Urban Syst.* **2008**, *32*, 331–342. [\[CrossRef\]](#)

11. Raubal, M.; Winter, S.; Teßmann, S.; Gaisbauer, C. Time geography for ad-hoc shared-ride trip planning in mobile geosensor networks. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 366–381. [[CrossRef](#)]
12. Miller, H.J. A Measurement Theory for Time Geography. *Geogr. Anal.* **2005**, *37*, 17–45. [[CrossRef](#)]
13. Trajcevski, G.; Wolfson, O.; Hinrichs, K.; Chamberlain, S. Managing uncertainty in moving objects databases. *ACM Trans. Database Syst.* **2004**, *29*, 463–507. [[CrossRef](#)]
14. Zipf, G. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*; Addison-Wesley Press: Boston, MA, USA, 1949.
15. Long, J.A. Modeling movement probabilities within heterogeneous spatial fields. *J. Spat. Inf. Sci.* **2018**, *16*, 85–116. [[CrossRef](#)]
16. Liao, F. Space-time prism bounds of activity programs: A goal-directed search in multi-state supernetworks. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 900–921. [[CrossRef](#)]
17. Demšar, U.; Long, J.A. Potential path volume (PPV): A geometric estimator for space use in 3D. *Mov. Ecol.* **2019**, *7*, 14. [[CrossRef](#)]
18. Kuijpers, B.; Technitis, G. Space-time prisms on a sphere with applications to long-distance movement. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 1980–2003. [[CrossRef](#)]
19. Argos. The Argos Tracking System. 2018. Available online: [www.argos-system.org](http://www.argos-system.org) (accessed on 1 July 2020).
20. Kuijpers, B.; Grimson, R.; Othman, W. An analytic solution to the alibi query in the space-time prisms model for moving object data. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 293–322. [[CrossRef](#)]
21. Giannotti, F.; Pedreschi, D., Eds. *Mobility, Data Mining and Privacy—Geographic Knowledge Discovery*; Springer: Berlin/Heidelberg, Germany, 2008.
22. Xie, Z.; Yan, J. Kernel Density Estimation of traffic accidents in a network space. *Comput. Environ. Urban Syst.* **2008**, *32*, 396–406. [[CrossRef](#)]
23. Winter, S. Towards a probabilistic time geography. In Proceedings of the GIS: ACM International Symposium on Advances in Geographic Information Systems, Seattle, WA, USA, 4–6 November 2009.
24. Winter, S.; Yin, Z.C. Directed movements in probabilistic time geography. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 1349–13665. [[CrossRef](#)]
25. Downs, J.A.; Horner, M.W.; Tucker, A.D. Time-geographic density estimation for home range analysis. *Ann. GIS* **2011**, *17*, 163–171. [[CrossRef](#)]
26. Song, Y.; Miller, H.J. Simulating visit probability distributions within planar space-time prisms. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 104–125. [[CrossRef](#)]
27. Horne, J.S.; Garton, E.O.; Krone, S.M.; Lewis, J.S. Analyzing animal movements using Brownian bridges. *Ecology* **2007**, *88*, 2354–2363. [[CrossRef](#)] [[PubMed](#)]
28. Rasmussen, C.E. Gaussian processes for machine learning. *Int. J. Neural Syst.* **2006**, *14*, 69–106.
29. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions*; Wiley: Hoboken, NJ, USA, 1994.
30. Burns, L. *Transportation, Temporal, and Spatial Components of Accessibility*; Lexington Books: Lexington, MA, USA, 1979.
31. Loraamm, R.W.; Downs, J.A.; Lamb, D.S. A time-geographic approach to quantifying wildlife-road interactions. *Trans. GIS* **2019**, *23*, 70–86. [[CrossRef](#)]
32. Downs, J.A.; Horner, M.W.; Hyzer, G.; Lamb, D.; Loraamm, R. Voxel-based probabilistic space-time prisms for analysing animal movements and habitat use. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 875–890. [[CrossRef](#)]
33. Hägerstrand, T. What about People in Regional Science? *Pap. Reg. Sci. Assoc.* **1970**, *24*, 7–21. [[CrossRef](#)]
34. Janelle, D.; Goodchild, M. Diurnal Patterns of Social Group Distributions in a Canadian City. *Econ. Geogr.* **1983**, *59*, 403–425. [[CrossRef](#)]
35. Miller, H. Modeling accessibility using space-time prism concepts within geographical information systems. *Int. J. Geogr. Inf. Syst.* **1991**, *5*, 287–301. [[CrossRef](#)]
36. Pfoser, D.; Jensen, C.S. Capturing the Uncertainty of Moving-Object Representations. In Proceedings of the Advances in Spatial Databases, 6th International Symposium, SSD'99, Hong Kong, China, 20–23 July 1999; Güting, R.H., Papadias, D., Lochovsky, F.H., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 1999; Volume 1651, pp. 111–132.
37. Egenhofer, M.J. Approximation of geospatial lifelines. In *SpadaGIS, Workshop on Spatial Data and Geographic Information Systems*; Bertino, E., Floriani, L.D., Eds.; University of Genova: Genoa, Italy, 2003.
38. Hornsby, K.; Egenhofer, M.J. Modeling Moving Objects over Multiple Granularities. *Ann. Math. Artif. Intell.* **2002**, *36*, 177–194. [[CrossRef](#)]

39. Wu, Y.H.; Miller, H. Computational tools for measuring space-time accessibility within transportation networks with dynamic flow. *J. Transp. Stat.* **2002**, *4*, 1–14.
40. Miller, H.J. Necessary space—Time conditions for human interaction. *Environ. Plan. B Plan. Des.* **2005**, *32*, 381–401. [[CrossRef](#)]
41. Othman, W. Uncertainty Management in Trajectory Databases. Ph.D. Thesis, Hasselt University, Hasselt, Belgium, 2009.
42. Timmermans, H.; Arentze, T.; Joh, C. Analysing space-time behaviour: New approaches to old problems. *Prog. Hum. Geogr.* **2002**, *26*, 175–190. [[CrossRef](#)]
43. Kwan, M.P. Gender and individual access to urban opportunities: A study using space-time measures. *Prof. Geogr.* **1999**, *51*, 211–227. [[CrossRef](#)]
44. Jacquez, G.M. Spatial analysis in epidemiology: Nascent science or a failure of GIS? *J. Geogr. Syst.* **2000**, *2*, 91–97. [[CrossRef](#)]
45. Jacquez, G.M.; Greiling, D.A.; Kaufmann, A.M. Design and implementation of a Space-Time Intelligence System for disease surveillance. *J. Geogr. Syst.* **2005**, *7*, 7–23. [[CrossRef](#)]
46. Löytönen, M. GIS, time geography and health. In *Location-Based Services*; GISDATA Series; Taylor & Francis: Abingdon, UK, 1998; pp. 97–110.
47. Neutens, T.; Witlox, F.; de Weghe, N.V.; Maeyer, P.D. Space-time opportunities for multiple agents: A constraint-based approach. *Int. J. Geogr. Inf. Sci.* **2007**, *21*, 1061–1076. [[CrossRef](#)]
48. Long, J.A.; Nelson, T.A.; Nathoo, F.S. Toward a kinetic-based probabilistic time geography. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 855–874. [[CrossRef](#)]
49. Long, J.A. Kinematic interpolation of movement data. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 854–868. [[CrossRef](#)]
50. Loraamm, R.W. Incorporating behavior into animal movement modeling: A constrained agent-based model for estimating visit probabilities in space-time prisms. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 1607–1627. [[CrossRef](#)]
51. Kuijpers, B.; Miller, H.J.; Othman, W. Kinetic prisms: Incorporating acceleration limits into space-time prisms. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 2164–2194. [[CrossRef](#)]
52. Shiryaev, A. *Probability*; Series: Graduate Texts in Mathematics; Springer: Berlin/Heidelberg, Germany, 1996; Volume 95.
53. Rice, J.A. *Mathematical Statistics and Data Analysis*, 3rd ed.; Duxbury Press: Belmont, CA, USA, 2006.
54. Slater, L. *Generalized Hypergeometric Functions*; Cambridge University Press: Cambridge, UK, 1966.
55. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*; 9th dover printing, 10th gpo printing ed.; Dover: New York, NY, USA, 1964.
56. Forbes, C.; Evans, M.; Hastings, N.; Peacock, B. *Statistical Distributions*, 4th ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2010.
57. OpenStreetMap. Available online: [www.openstreetmap.org](http://www.openstreetmap.org) (accessed on 1 July 2020).
58. Downs, J.A.; Lamb, D.; Hyzer, G.; Loraamm, R.; Smith, Z.J.; O’Neal, B.M. Quantifying spatio-temporal interactions of animals using probabilistic space-time prisms. *Appl. Geogr.* **2014**, *55*, 1–8. [[CrossRef](#)]
59. Furstenburg, D. Focus on the African Buffalo (*Syncerus caffer*). *S A Hunt.* **2010**, *05040*, 46–49.
60. Cross, P.; Bowers, J.; Hay, C.; Wolhuter, J.; Buss, P.; Hofmeyr, M.; du, Toit, J.; Getz, W. Movebank: Kruger African Buffalo, GPS tracking, South Africa. Available online: [www.movebank.org/cms/webapp?gwt\\_fragment=page=studies,path=study1764627](http://www.movebank.org/cms/webapp?gwt_fragment=page=studies,path=study1764627) (accessed on 1 July 2020).

