Discrete Partially Ordered Sets in Physics

Geoffrey Hemion

Contents

Chapter 1

The Two Slit Experiment

1.1 Discrete versus continuous models

It is sometimes said that the essential idea of quantum mechanics is best illustrated by the two slit thought experiment. As in an old fashioned cathode ray tube we have a source of electrons, call it S, an obstructing screen some distance from S with two parallel slits which are close together, call the slits A and B, and a fluorescent screen, call it F, beyond the two slits, which lights up each time an electron hits it. The classical physicist would think that a diffuse, Gaussian sort of glowing pattern should appear on the fluorescent screen, regardless of whether there are one, two, or more slits in the obstructing screen. After all, given that electrons are like little billiard balls, then they go through one or the other slit, and in doing so they possibly bounce about on the billiard-like particles making up the walls of the slit, deflecting them in some way or another to different angles onto the fluorescent screen, independent of whichever slit they happen to pass through. Uncertainty arises from the fact that it is, as a practical matter, impossible to know the details of the movements of all the billiard-like particles in the experiment. In principle, if we knew all of these details then by following the equations of physics we could calculate the exact outcome each time.

But as we know, quantum mechanics confounds this simple picture. Rather than having a Gaussian type of pattern, instead we see a wave-like pattern, as if the electron has suddenly stopped being a billiard ball and instead become something like waves of light progressing through the aether, passing through the two slits simultaneously and then producing waves of interference in the space beyond the obstructing screen. For the classical physicist, this strange behavior of the electron, being both a small billiard ball and an aetherial wave, was a mystery.

Perhaps the mystery can be resolved if we think about the experiment in a different way, using a different mathematical model which involves neither billiard balls nor aetherical waves. For this model we will retain the picture of space being three dimensional Euclidean space and time being measured using the real number line, giving a four dimensional space with time.

According to the billiard ball model of the two slit experiment, the electron travels in a straight line from the source S to either A or B, then after going through either of the slits, it travels to some particular point P on the fluorescent screen F, again following a straight line. Considered in the four dimensions of space and time, the electron follows a path consisting of two straight line segments. Our model is similar to this, but with the path of the electron now being a discrete set of points, equally

spaced at intervals representing the wavelength of the supposed aetherial electron wave. These discrete paths are taken to be straight from S to A, or to B, and then straight to the point of landing P. In addition to this, all of the rest of the experiment – the cathode source S, the obstructing screen, the fluorescent screen, and in fact the whole universe surrounding everything – is taken to consist of similarly discrete particles: discrete paths of points in space and time.

But now we must think about the role of probability in our model. As we have seen, the classical physicist imagines that probability arises through a lack of perfect knowledge of all the details of the experiment. In those terms the uncertainty introduced by this discreteness would, indeed make the calculations of classical physics even more complicated, but in principle it is the same classical model. It is based on the idea that if we put some sort of box around the experiment in space and time, then if we knew everything about the "outside world" outside the box, it would be possible to use the laws of physics to calculate exactly what has happened within the box. Or put anther way, given a fixed configuration of the universe outside the box, there is only one, fixed, possible configuration of the experiment within the box.

Now surely the modern physicist, unlike his classical colleague, is prepared to be open to other possibilities. If quantum mechanics implies that uncertainty exists as an essential part of the physical world then we might imagine that even if a fixed configuration of the universe is given outside our box, within the box various different configurations might be possible. If this is the case then we arrive at a different method of defining probabilities.

In the case of the two slit experiment, let us take two possible points, P and P', for the electron to land on the fluorescent screen. P is a point of constructive interference and P' is a point of destructive interference. Using the wave picture, the two possible paths for the electron, going through either A or B to the point P, following straight line segments, differ in length by an exact multiple of the wavelength. And then the lengths of the straight line paths from S to A or B, and then to P', differ by a multiple of the wavelength plus half a wavelength. Therefore we have two different experiments: one leading to the result P and the other leading to the result P'. Let us call these the experiments E and E'.

So what are the relative probabilities of the two experiments E and E'? Our electron comes from the outside universe into the box, passes through it, and emerges on the other side back out into the outside universe. The outside universe is fixed, while the events within the box are unknown to us, allowing various possibilities. In particular the discrete path of the electron outside of the box, when it enters and when it leaves, is given.

But now it is obvious that E will be more probable than E'. Given a particular run of the experiment, we have a discrete path involving straight segments from S to either A or B and then to P, finally emerging from the box correctly synchronized to connect up with its path into the outside universe. But then, if one possible path was through A, another possible path would be through B, and again the electron would be correctly synchronized to connect up with its path into the outside universe. Thus we have two possible universes which contain this experiment. Contrast this with the experiment E'. There, if say the electron goes through slit A and connects properly with itself when emerging from the box then the alternative variation through B is blocked. The electron will not have the right synchronization to emerge into the fixed outside universe. There is only one possible variation and so E' is only half as probable as is E.

Of course all the details of these fixed outside universes are not important. The rule for calculating such local probabilities is that we assume that all possible outside

universes are equally likely; then each of these is multiplied by the number of local variations with which they are compatible. Thus, in reality, it is sufficient to simply take all possible boundaries of the boxes and examine how many different variations of the configurations within the boxes are compatible with these boundary values.[1]

1.2 The wave function

The result of the two slit experiment is not just a simple wave. The fluorescent screen is lit up more strongly in the middle, near where the slits are. Looking to the right or left it grows dimmer, as if a Gaussian curve is superimposed upon the wave, since it is unlikely that the electrons are deflected way away from a more direct path, and even the points of destructive interference are not completely dim. So we will imagine that the quantum mechanical part of the experiment refers just to the results which depend exclusively on the phase differences of the various possible paths. For the sake of simplicity, all other factors will be ignored.

In the middle of the fluorescent screen, directly in front of the two slits, with the maximum constructive interference, the intensity is strongest. Then moving to the right or left we reach the first regions of destructive interference with reduced intensity. By subtracting a suitable normalizing constant, let us say that a point of constructive interference is given the value $+1$ and a point of destructive interference is -1. We would like to show that the brightness, representing the quantum mechanical part of the relative probabilities, is given by the cosine function.

Let us say that we have one possible discrete path in an experiment, call it γ_1, and there is another, different discrete path, γ_2, which may or may not be fully compatible with γ_1 in the sense that the discrete phases might not completely line up going into and out of the experiment. But even if they do not line up perfectly, still, perhaps when taking a slightly more expansive view of the experiment it might be that an expanded version of γ_2 interacts with other particles and does, in the end, allow itself to line up with an expanded version of γ_1. Be that as it may, we will say that if the phases of γ_1 and γ_2 line up perfectly, then the "influence" of the existence of γ_1 on the probability of γ_2 also occurring has the value $+1$. And if they are perfectly out of phase then the influence has the value -1. For phase differences between those two extremes, the influence function is something between $+1$ and -1.

So let u be a real-valued function which assigns a value to the difference $l_1 - l_2$ of the path lengths of γ_1 and γ_2, representing the influence of γ_2 on γ_1 when calculating the quantum mechanical probability. Then we will have u being a symmetric function:

$$u(\theta) = u(-\theta).$$

And in order to tie in with more usual ways of thinking, we will assume that the discrete path lengths have been somehow normalized so that the distance between

[1] It might be objected that the experiment E' also has two possible scenarios, namely the particle could go through slit A and connect up with one possible outside universe, and also it could go through slit B and connect up with some other *completely different* outside universe. But if we are allowed to consider two completely different outside universes for the experiment E' then surely we should also be allowed to do the same thing for the experiment E. Taking some different synchronization (perhaps thought of with respect to the discrete particles which make up the slits A and B) and an appropriate outside universe allowing this synchronization, we then again have two possible paths for the particle leading to E with respect to this new outside universe.

More generally we can consider all boxes leading to the result E and all boxes leading to the result E'. The assumption is that there are more or less equal numbers of outside universes compatible with these various boxes. Then the number of possible universes (including the boxes) is multiplied by 2 for the result E and only by 1 for the result E' Therefore E is more probable.

the points on the paths is 2π. Thus we have

$$u(k \cdot 2\pi) = 1$$

and

$$u(k \cdot 2\pi + \pi) = -1,$$

for all integer valued k.

We would like to prove that $u = \cos$. One way to do this, following Gudder[2], is to assume that for any finite set of numbers $\theta_1, \ldots, \theta_n$ which are such that

$$\sum_{i=1}^{n} u(\theta_i) = 0$$

then we also have

$$\sum_{i=1}^{n} u(\theta_i + \tau) + \sum_{i=1}^{n} u(\theta_i - \tau) = 0,$$

for all real numbers τ. Functions which satisfy this condition will be called "causal".

Theorem. *If u is causal, continuous and has a zero, then there exists a $k > 0$ such that $u(a) = u(0)\cos(a \cdot k)$ for all $a \in \mathbb{R}$.*

Proof. We will prove this theorem in five steps, closely following Gudder's proof.

(i) Can we have $u(0) = 0$? If so, then the causality property of u implies that $u(a+0)+u(a-0) = 2u(a) = 0 \Rightarrow u(a) = 0$, for all $a \in R$, and the theorem is trivially true. Therefore we will assume that $u(0) \neq 0$. Without loss of generality we may assume that $u(0) = 1$.

(ii) Let d be the smallest positive number such that $u(d) = 0$. (The fact that d exits follows from the assumption that u has a zero.) We have $u(a+d)+u(a-d) = 0$ for all a, in particular for $a = d$. Therefore $u(0) + u(2d) = 0$, or $u(2d) = -u(0) = -1$. More generally we have

$$u((a+d)+d) + u((a+d)-d) = u(a+2d) + u(a) = 0,$$

that is, $u(a+2d) = -u(a)$, for all $a \in \mathbb{R}$.

(iii) Assume that the number b is such that $u(b)$ is rational, with $u(b) = s/t$, say, where s and t are integers (we can assume that t is positive.) We will now prove that $u(a+b) + u(a-b) = 2u(a)u(b)$ for all $a \in R$. For let $\{a_1, \ldots, a_{s+t}\}$ be the following set of real numbers. $c_1 = \cdots = c_s = 2d$ and $c_{s+1} = \cdots = c_{s+t} = b$. We then have

$$\begin{aligned}
\sum_{i=1}^{s+t} u(c_i) &= -s + t \cdot u(b) \quad = \quad 0 \quad \Rightarrow \\
0 &= \sum_{i=1}^{s+t} \left(u(a+c_i) + u(a-c_i)\right) \\
&= s\left[u(a+2d) + u(a-2d)\right] + t\left[u(a+b) + u(a-b)\right] \\
&= -2s \cdot u(a) + t \cdot \left[u(a+b) + u(a-b)\right]
\end{aligned}$$

Thus

$$u(a+b) + u(a-b) = \frac{2s}{t} u(a) = 2u(a)u(b)$$

[2]Amplitudes and the Universal Influence Function, Journal of Mathematical Physics 32, 2106 (1991)

and therefore we have established the result for all b with the property that $u(b)$ is rational. Since the function u was assumed to be continuous, we conclude easily that the result holds for all real b.

(iv) Set $a = b$. Then we have

$$\begin{aligned} u(2b) + u(0) &= 2u(b)^2 \quad \Rightarrow \\ u(2b) + 1 &= 2u(b)^2 \quad \Rightarrow \\ u(b) &= \left(\frac{u(2b)+1}{2}\right)^{\frac{1}{2}} \end{aligned}$$

(v) We may now finally prove that $u(a) = \cos(a \cdot \pi/2d)$ for all real numbers a of the form

$$a = \frac{n \cdot d}{2^m}$$

where n and m are integers, with $m \geq 1$, using induction on m. According to the steps (i) and (ii), the result is true for the case $m = 1$. Assume therefore that $m > 1$. We have

$$\begin{aligned} u(a) &= \left(\frac{u(2a)+1}{2}\right)^{\frac{1}{2}} = \left(\frac{\cos(2a)+1}{2}\right)^{\frac{1}{2}} \\ &= \cos(a) \end{aligned}$$

The last equation follows from the inductive hypothesis and elementary trigonometry.

Finally, one need only notice that the numbers of the form $n \cdot d/2^m$ are dense in the set of all real numbers to conclude that $u(a) = \cos(a)$ for all $a \in R$. □

In order to show that our influence function u is causal we will consider a modified version of the two slit experiment. We begin with the simple case that we have some phase difference θ which is such that $u(\theta) = 0$. We must show that

$$u(\theta + \tau) + u(\theta - \tau) = 0,$$

for all possible τ.

In this case the modified experiment includes a segment before S and a segment after P, both of which consist of further two slit experiments. The segment before S has the electron starting at the new source S', traveling through one of two slits and then passing through S into the main experiment. These two new slits are such that the path length through the left hand slit is τ greater than the path length through the right hand slit. Then the segment after P consists of another two slit experiment such that the path length through the left hand slit is $-\tau$ compared to the path length through the right hand slit.

Let γ be the path from S' to P' along the right hand path to S then along the given path in the main experiment and finally along the right hand path to P'. We consider the influence the existence this path has on the probability that paths through the alternative route through the main experiment are taken. For simplicity, let us assume that the lengths of the right hand paths are both zero modulo 2π. That is, they are multiples of the discrete distance between adjacent elements along the paths. Then, following the various possible paths through the modified experiment, the total influence will be:

$$u(0+\theta+0) + u(\tau+\theta+0) + u(0+\theta-\tau) + u(\tau+\theta-\tau) = 2u(\theta) + u(\theta+\tau) + u(\theta-\tau).$$

On the other hand, we can consider the two possible paths from S' to S, then the given path through the main experiment, and then the two possible paths from P to P' as giving four different experiments with fixed paths in and out of the main experiment whose purpose is to compare possible varied paths just within the main experiment. In each case the influence of the alternative path will then be $u(\theta) = 0$, since the phase difference between the given and the alternative path in the main experiment is θ. Therefore

$$2u(\theta) + u(\theta + \tau) + u(\theta - \tau) = u(\theta + \tau) + u(\theta - \tau) = 0,$$

showing that the rule for causality holds in this case.

More generally, let $\theta_1 \ldots, \theta_n$ be such that

$$\sum_{i=1}^{n} u(\theta_i) = 0.$$

We modify the main experiment, adding in additional slits so that we have the slit representing the given path, and then there are n further slits, such that for the path going through the i-th slit, the phase difference to the given path is θ_i. With this, the argument is then analogous to the case where $n = 1$.

Of course there is much more to physics than the two slit experiment. Nevertheless, it can be used as a model for explaining why the method of Feynman diagrams successfully describes quantum electrodynamics. According to Feynman's description, the diagrams consist of straight line segments connecting points of interaction. Each such segment is assigned a value proportional to an expression of the form $\exp(-iS(\xi))$, where S is a certain real function of the path ξ. (In our simple treatment of the two slit experiment we have taken this function to be the path length.) Generalizing things, let us say that a given experiment allows a certain set Ω of possible configurations leading to a given result. We assume that Ω is finite, thus avoiding the divergent integrals in the usual continuous description. Taking our influence function between two configurations ξ and $\chi \in \Omega$ to be

$$u(S(\xi) - S(\chi)) = \cos(S(\xi) - S(\chi)),$$

we can then say that the probability of Ω is proportional to[3]

$$\begin{aligned}
\sum_{\xi,\chi\in\Omega} \cos(S(\xi)-S(\chi)) &= \frac{1}{2}\sum_{\xi,\chi\in\Omega}\{e^{i(S(\xi)-S(\chi))}+e^{-i(S(\xi)-S(\chi))}\} \\
&= \frac{1}{2}\sum_{\xi,\chi\in\Omega}\{e^{i(S(\xi)-S(\chi))}+e^{i(S(\chi)-S(\xi))}\} \\
&= \sum_{\xi,\chi\in\Omega} e^{i(S(\xi)-S(\chi))} \\
&= (\sum_{\xi\in\Omega} e^{iS(\xi)})(\sum_{\chi\in\Omega} e^{-iS(\chi)}) \\
&= |\sum_{\xi\in\Omega} e^{iS(\xi)}|^2
\end{aligned}$$

In other words, the probabilities can be calculated by summing the complex amplitudes over all possible configurations of the underlying model, and then taking the square of the absolute value. This is the usual prescription for dealing with the amplitude function in quantum mechanics.

[3]One way to think about the double sum here would be to assume that, as would be the case in "classical" physics, each possible path through an experiment is compatible with a *unique*, ideal outside universe. Then the second sum represents all the further, not quite ideal, variations within the experiment for that given outside universe. More generally the assumption is that all the possible outside universes which are compatible with paths within the experiment can be divided up into approximately equally large equivalence classes, with each such class being associated with a different path such that the outside universes in a given equivalence class all contain the path with which they are associated in a relatively ideal way.

Chapter 2

Partially Ordered Sets

2.1 Philosophical thoughts and a definition

Could it be that it is possible to describe the physical world in terms of some *ultimately simple* principle?

Theoretical physics, as it is known today, is certainly not simple. Everything is based on "field theories". Various kinds of continuous functions on differential manifolds. Even topological manifolds with complicated connectivity properties are brought in. The basis of analysis is the system of real numbers. How many axioms are needed to define the real numbers? What are the additional axioms for describing vector spaces? Then Euclidean spaces; complex numbers; tensors; topology; manifolds; limit theorems, and so on and so forth. Does anybody seriously believe that this is the basis for some ultimate theory of the physical world? But each step in the accumulation of this vast system was undoubtedly plausible at the time it was added.

After all, looking about us we see things floating through the air, seemingly completely smoothly. Various schools of philosophy in ancient Greece concluded that the world consisted of something called "space", complete with particles, or "atoms", traveling about within this space. This is what we see when looking about at the world around us. But to use a more modern philosophical analogy, what if the world is not really what we see? I am writing this text on a computer and watching the words appear on the screen as I type. But this is only the interface to a deeper reality, making things comprehensible to my eyes and my brain. The reality is that my typing causes complex patterns of electrical activities in microscopically small, yet immensely complicated circuits in the computer which I could never comprehend in detail, perhaps later resulting in electromagnetic interactions, again, in detail, incomprehensible, in a vast network connecting an unknown number of further computers around the world. Yet this interface which I have before me, despite the fact that it is only an illusion, allows me to function in some rational way.

Is the concept of space, and of particles traveling about within space, similarly an illusion, an interface whose purpose is to allow us to function in the real world? Something which should not be taken as being the basis of reality.

It is certainly not the intention here to delve further into meaningless philosophical speculation. Instead, let us try to imagine the simplest possible mathematical structure, which is not so simple as to be totally trivial, and yet which might allow a sufficient degree of complexity in order to form the basis for a description of the physical world.

For example we have the real number line. This is defined to be a field (in the sense of pure mathematics). Thus we have the axioms for addition and multiplication, associativity and commutitivity. It is an ordered field with the order respecting the arithmetical operations. And then it is Dedekind-complete. In addition to being overly complicated in its definition, it has the additional drawback of having a rigid, one dimensional structure. It is fixed. The idea that the world *is* the real number line is simply absurd.

Other, still simpler mathematical frameworks present themselves. For example, could it be that the world is a group? Recall that a group is a set G containing various elements. There is a binary operation which is associative. There is an identity element, and for each element there is an inverse element. Clearly the framework of groups is simpler and more comprehensive than that of real numbers. In fact the concept of groups is used in the the definition of real numbers. There have been many, very deep mathematical investigations into group theory. Perhaps an argument could be made to support the proposition that the world is, in its basic structure, a group. But I cannot see it.

Instead, we will base everything on the concept of partially ordered sets. A partially ordered set is a pair X, together with a subset $\preceq \subset X \times X$ of the Cartesian product of X with itself which fulfills the conditions:

1. $(a, a) \in \preceq$ for all $a \in X$, (Reflexivity)
2. if $(a, b) \in \preceq$ and $(b, a) \in \preceq$ then $a = b$, (Antisymmetry)
3. if $(a, b) \in \preceq$ and $(b, c) \in \preceq$ then $(a, c) \in \preceq$, (Transitivity).

A more familiar way to describe these conditions is to say that the notation $(a, b) \in \preceq$ means $a \leq b$, or "a is less than or equal to b". Using this notation, we can say that a partially ordered set is a pair $(X, \leq)$ satisfying our conditions. For brevity it is usual to say "poset" rather than "partially ordered set".

There is a natural partial order already given in most of the usual theories of physics. For let x and y be two points of "space-time". Then if we take $x \leq y$ to mean that y is in the light cone above x, we obtain a partial ordering of some sort of imagined space-time manifold. Or thinking in other terms, we might say that $x \leq y$ means that it is possible to send a message from x to y, or that y is a "consequence of" x,

But such *continuous* posets are surely too complicated to qualify as being things of ultimate simplicity. Going to the opposite extreme, taking X to be the empty set, leaves us with nothing, which is certainly not an appropriate model for the physical world. Similarly *finite* posets, although they have been the subject of much mathematical investigation, seem to be too restricted for our purposes. If we have a finite poset then we have special elements which are *extreme*. That is to say they are *maximal* or *minimal*: there are no further elements which are greater than, or no further elements which are less than the given element. Elements which are not extreme are *interior* elements. Surely the existence of extreme elements adds a level of complexity which we can do away with.

Some additional notation is called for here. Let X be a poset and let $a, b \in X$. If either $a \leq b$ or $b \leq a$ then a and b are *related*, written $a \perp b$. If a and b are not related then we write $a \| b$. Of course $a < b$ means $a \leq b$ and $a \neq b$. If $a \leq b$ then we write $a \between b$ to denote the set of elements between a and b. That is

$$a \between b = \{c \in X : a \leq c, \text{ and } c \leq b\}.$$

Given $a \in X$ then $a_\downarrow$ is the set of elements less than a, and $a_{\Downarrow}$ is the set of elements less than or equal to a. Which is to say

$$a_\downarrow = \{c \in X : c < a\}, \text{ and } a_{\Downarrow} = \{c \in X : c \leq a\}.$$

The definitions of $a_\uparrow$ and $a_{\Uparrow}$ are analogous, referring to elements greater than a.

Now in order to avoid the unnecessary complexity of continuity, let us specify that only non-empty, *discrete* posets containing no extreme elements are to be allowed. What do we mean by discrete? Perhaps the usual idea would be to say that it means $a \between b$ is finite for all $a < b$. This is the idea used in the definition of *causal sets*. But we will choose a somewhat more restrictive definition which could be termed "strongly discrete". Namely we will say that X is *discrete* if $a_\downarrow \setminus b_\downarrow$ is finite for all a and b in X. Here "\" denotes the set difference, the set of all elements in $a_\downarrow$ which are not in $b_\downarrow$.

The poset is called *connected* if for any two elements a, $b \in X$, there exists a (finite) sequence of elements of X, starting with a and ending with b, such that adjacent pairs of elements in the sequence are always related. It is no restriction to only allow connected posets since otherwise we could deal with the different connected components separately.

We will also specify a condition which is analogous to the axiom of extensionality in set theory. A poset X will be called *extensional* if for all a, $b \in X$ with $a \neq b$ we have both $a_\downarrow \neq b_\downarrow$ and $a_{\Downarrow} \neq b_\downarrow$.

The poset X will be called *confluent below* if for any two elements a, $b \in X$ we have $a_\downarrow \cap b_\downarrow \neq \emptyset$.

Then finally, the poset X will be called *upwardly separating* if for any two elements a and b with $b \not\geq a$ we have $b_\uparrow \setminus a_\uparrow \neq \emptyset$. (Note that the condition that a poset is upwardly separating implies by itself that the poset contains no maximal elements.)

Definition 1. *Let us denote by $\mathfrak{W}$ the class of all non-empty, discrete, upwardly separating, confluent below, connected posets X, fulfilling the condition of extensionality, such that all elements of X are interior.*

The aim here will be to argue that the physical world can best be described in terms of some particular poset $X \in \mathfrak{W}$. Since we only have limited information about the details of everything in the world, we cannot know precisely which poset in $\mathfrak{W}$ is the one which describes the actual world. We must consider all possible posets which are compatible with the information we have. And therefore we will be dealing with a probability space, based upon the class $\mathfrak{W}$ of posets.

Are the properties which we are using to define $\mathfrak{W}$ in some sense "natural"? Or would it be better to use some other properties, giving a different class of posets — or indeed, would it be better to use some entirely different mathematical structure all together?

To begin it might be useful to compare our framework for defining $\mathfrak{W}$ with the usual theory of sets. Thinking in these terms, the elements of a poset $X \in \mathfrak{W}$ would be themselves sets, and the ordering of X would be given by set inclusion.

The property of extensionality is one of Zermelo-Fraenkel's axioms. However the requirements that X be confluent below, and particularly that there be no extreme elements, are a decided departure from Zermelo-Fraenkel set theory. According to Zermelo-Fraenkel, the possibility of an infinitely descending chain of set inclusions is explicitly excluded.

Another way to think about our framework would be to compare it with the idea of cause and effect. Or in logic, the idea of assumptions leading to conclusions. If

a and $b \in X$ are two events, then we will write $a < b$ to mean that a is one of the causes of b. Or b is one of the effects of a. Then we might contemplate the question of whether or not the conditions defining $\mathfrak{W}$ are reasonable ones for describing cause and effect.

2.2 Chains and anti-chains

Within the theory of posets, the ideas of chains and anti-chains are important. A *chain* $C \subset X$ is a totally ordered subset. That is, if a, $b \in C$ then $a \perp b$. An anti-chain is a subset $A \subset X$ such that for any two elements a, $b \in A$ with $a \neq b$, we have $a \| b$.

A chain is *maximal* if it cannot be properly contained in another chain. Similarly a maximal anti-chain cannot be properly contained in another anti-chain. Obviously, a maximal anti-chain $Y \subset X$ is such that every element of X is either in Y, or at least related to an element of Y.

Since X is extensional, we cannot have a maximal anti-chain consisting of just a single element. But much more than this, we have the following theorem.

Theorem 2.1. *For every poset in $\mathfrak{W}$, all maximal chains and maximal anti-chains are infinite.*

Proof. The fact that maximal chains are infinite follows from the condition that all elements are interior elements. To see that maximal anti-chains are also infinite, assume to the contrary that $A \subset X$ is a finite maximal anti-chain. Since X is confluent below, there exists an element a which is less than all elements of A. Choose an element $b \in X$ with $b \| a$. Such an element b must exist since we can choose some $d < a$ with $d_\uparrow \cap a_\downarrow = \emptyset$ and then observe that the set $d_\uparrow \setminus a_\uparrow \neq \emptyset$, since X is upwardly separating. Similarly we have $b_\uparrow \setminus a_\uparrow \neq \emptyset$. In fact though, the set $b_\uparrow \setminus a_\uparrow$ must be finite. To see this, begin by noting that there can only be finitely many elements of X which are greater than b, yet less than an element of A. On the other hand, if $c \in b_\uparrow \setminus a_\uparrow$ is not less than any of the elements of A, then we must have $c \| A$. That is c is unrelated to all the elements of A, contradicting the fact that A is a maximal anti-chain.

Thus $b_\uparrow \setminus a_\uparrow$ must be finite but non-empty and so we can choose a maximal element $h \in b_\uparrow \setminus a_\uparrow$. But then we must have $h_\uparrow \setminus a_\uparrow = \emptyset$, which contradicts the condition that X is upwardly separating. □

Later on, we will consider what might be called "generalized" chains, as being models for the elementary particles of physics.

2.3 Positions

Given any poset $(X, \leq)$ (not necessarily in $\mathfrak{W}$), we can define the set of *positions* within X, as follows.

Definition 2. *A position $\mathcal{P} \subset X$ consists of a pair of non-empty subsets U, $V \subset X$, such that $U \leq V$ (that is, $u \leq v$ for all $u \in U$ and $v \in V$) and such that the pair is maximal in the sense that if U is properly contained in U', then we cannot have $U' \leq V$, and also if V is properly contained in V' then we cannot have $U \leq V'$. We also write $\mathcal{P}_\downarrow$ for U and $\mathcal{P}_\uparrow$ for V.*

Given any element $a \in X$, then the pair $(a_{\Downarrow}, a_{\Uparrow})$ forms a position in X. We will call such positions *elementary positions*. The set $\Omega(X)$ of all possible positions in X is itself a poset in a natural way. It contains X, but in general it is much larger than X. One could say that $\Omega(X)$ is the *completion* of X. If X is equal to its completion, then we say that X is *complete*. Within the theory of finite posets, if we add in a single minimal element and a single maximal element, then the completion is the *Macneille completion*, which is a lattice (in the sense of mathematical combinatorics).

Any position $\mathcal{P} = (U, V)$ is determined by its lower and upper sets U and V. After all that is the *definition* of the position. But it may be possible to find two subsets $U_* \subset U$ and $V_* \subset V$ such that $\mathcal{P}$ is the *only* position lying between U_* and V_*. In this case we can say that $\mathcal{P}$ is determined by the pair (U_*, V_*).

Definition 3. *Let U_*, $V_* \subset X$ be two subsets, such that $U_* \leq V_*$. If there is only one position $\mathcal{P}$ in X such that $U_* \leq \mathcal{P} \leq V_*$, then we will say that $\mathcal{P}$ is* determined by *the pair (U_*, V_*). The pair will be called* minimal*, if there is no smaller pair (Y, Z) with $Y \subset U_*$ and $Z \subset V_*$ which also determines $\mathcal{P}$. When considering pairs of subsets which determine a given position, we will usually assume that the pair is minimal.*

Now it is obvious that each elementary position is determined by just one single element, namely the element which the position represents. Furthermore, if $\mathcal{P}$ is a non-elementary position, determined by the minimal pair (Y, Z), then both Y and Z must have at least two elements.

Another way to look at these things is the following. Let $U_* \subset X$ be some subset such that $U_{*\uparrow} \neq \emptyset$. Then take $V = U_{*\uparrow}$, and $U = V_{\downarrow}$. If we assume that $U_{*\uparrow}$ is not the upper set of some element of X, then the pair (U, V) is a position $\mathcal{P}$ in X. Therefore, given that $\mathcal{P}$ is determined by some pair (U_*, V_*), then we can also say that $\mathcal{P}$ is determined by the lower set U_* alone, following this procedure. Analogously, a position can be determined by an upper set.

At this stage, it is useful to consider a further idea.

Definition 4. *Let $(X, \leq)$ be a poset (again, not necessarily in $\mathfrak{W}$), and let $\mathcal{P}$ be a position in X. We will say that an element $a \in X$ is* associated with *$\mathcal{P}$ if $\mathcal{P} \setminus \{a\}$ is* not *a position in $X \setminus \{a\}$. If an elementary position is associated with itself (that is, with the element generating the position), then we will say that the element is an* essential *element. Otherwise, the element is non-essential; it can simply be removed without affecting the set of positions of X.*

We now confine our attention to posets in our class $\mathfrak{W}$.

Theorem 2.2. *Let $a \in X$ be associated with the non-elementary position $\mathcal{P} \subset X$. Then $a < \mathcal{P}$. That is $a \in \mathcal{P}_{\downarrow}$.*

Proof. Let $\mathcal{P} = U \cup V$ with $U \leq V$. If $a \notin \mathcal{P}$ then $\mathcal{P} \setminus \{a\} = \mathcal{P}$. Since a is associated with $\mathcal{P}$, it must be that the pair (U, V) is not maximal in $X \setminus \{a\}$. But that implies that (U, V) is not maximal in X, which is a contradiction.

If $a \geq \mathcal{P}$ (that is, $a \in V$, the upper set of the position), then since $\mathcal{P} \setminus \{a\}$ is not a position in $X \setminus \{a\}$, it must be that the pair $(U, V \setminus \{a\})$ is not maximal in $X \setminus \{a\}$. That is, there must be an element $b < V \setminus \{a\}$, such that $b \notin U$, and so $b \| a$. But since X is upwardly separating, there exists some $c \in a_{\uparrow} \setminus b_{\uparrow}$. Since a is in the upper set of $\mathcal{P}$, we must have c also being in the upper set. i.e. $c \in V \setminus \{a\}$. However, this contradicts the fact that $b < V \setminus \{a\}$. □

Theorem 2.3. *Assume that the non-elementary position $\mathcal{P}$ is determined by the minimal pair (U_*, V_*), where $U_* < V_*$. Assume furthermore that A is the set of all*

elements of X which are associated with $\mathcal{P}$ (and therefore $A \subset \mathcal{P}$). Then we have $A \subset U_*$.

Proof. Let $a \in A$. Since $\mathcal{P}$ is associated with a, we must have another position $\mathcal{R}$ in X with $\mathcal{P}_\downarrow = \mathcal{R}_\downarrow \cup \{a\}$. If $a \notin U_*$ then we would have both $\mathcal{P}$ and also $\mathcal{R}$ being between U_* and V_* so that $\mathcal{P}$ is not determined by the pair (U_*, V_*). This is a contradiction. □

We have shown that $A \subset U_*$. Nevertheless our given formulation of the theory does not rule out the possibility that $U_* \backslash A \neq \emptyset$, but we expect this to be the exception. The elements of $U_* \setminus A \neq \emptyset$ could be thought of as being associated with the position in a more generalized sense. In any case the position $\mathcal{P}$ is determined by the set U_*. And for the sake of simplicity in further arguments where it is of no consequence, we will generally assume that positions are just determined by the elements with which they are associated.

2.4 General properties of the physical world

1. The past and the future are different from one another. The future develops out of the past. Similarly in $\mathfrak{W}$, given that the ordering of a poset $X \in \mathfrak{W}$ represents time, then we have the result that all non-elementary positions $\mathcal{P}$ are determined by elements of X which come *before* $\mathcal{P}$ in time.

2. The universe appears to be expanding. That is to say, the light from distant galaxies appears to be red-shifted. Put another way, it appears that the measure of time in the earlier universe is longer than it is now; time appears to be speeding up. Similarly in $\mathfrak{W}$, if a, $b \in X$ are two related elements in a poset in $\mathfrak{W}$ with say $a < b$, then a certain measure of time could be obtained by taking the number of elements in $b_\downarrow \setminus a_\downarrow$. This is reflected in the fact that X is discrete, yet upwardly separating.

3. The material of the world appears as particles, many of which persist for long periods of time. Again, $\mathfrak{W}$ certainly contains posets whose elements are not arranged in tight chains. But perhaps there is some probabilistic effect which would lead us to the conclusion that *most* such posets are composed of collections of chains.

Chapter 3

The Geometry of Posets

3.1 Dimension

For finite posets the idea of dimension is defined in the following way[1]. Given a poset X with its partial ordering $\preceq$, it is *linearly ordered* if it is simply a chain. Otherwise there must exist at least two elements a, $b \in X$ which are unrelated: $a \| b$, or in other words, neither $(a, b) \in \preceq$ nor $(b, a) \in \preceq$. But then we can simply add in either the pair (a, b) to $\preceq$ to create the new partial order $\preceq_1$ or else we can add in the pair (b, a) to $\preceq$ to create the new partial order $\preceq_2$. It is a simple matter to confirm that the three rules for partial orderings will still apply. After a finite number of such steps, we obtain a linear order $\preceq_L$ which contains the original order $\preceq$. (In fact one sees that the set of all possible partial orders on a finite set X is itself a partially ordered set, with the ordering given by set inclusion.) $\preceq_L$ is called a *linear extension* of $\preceq$.

A set of linear extensions $\{\preceq_{L_1}, \ldots, \preceq_{L_n}\}$ of $\preceq$ such that their intersection is $\preceq$ is called a *realizer* of $\preceq$, that is

$$\bigcap_{i=1}^{n} \preceq_{L_i} = \preceq .$$

If the poset has a realizer then the dimension of the poset is defined to be the least possible number of linear extensions in a realizer. It is not difficult to see that if a poset $(X, \leq)$ has the dimension n, then it can be embedded in an order preserving way in n-dimensional Euclidean space $\mathbb{R}^n$, where the partial ordering of $\mathbb{R}^n$ is given by the rule: $(x_1, \ldots, x_n) \leq (y_1, \ldots, y_n)$ if $x_i \leq y_i$ for all $i = 1, \ldots, n$.

I have discussed a somewhat generalized version of this definition, perhaps more suitable for physics, in a recent paper.[2] Nevertheless, it is unclear whether or not typical posets $X \in \mathfrak{W}$ can be expected to be 4-dimensional in any sense at all. And certainly $\mathfrak{W}$ contains posets of arbitrarily large dimension. In fact, since all posets in $\mathfrak{W}$ are infinite, we would expect them to have no fixed dimension at all. At most we can expect that 4-dimensionality might arise through some sort of probabilistic effect, at least at a local level.[3]

The interesting thing here is that dimension – a property of geometry – is defined independently of any of the structures (vector spaces, differential manifolds) which

[1] The book, "Combinatorics and Partially Ordered Sets", by William T. Trotter, contains many further results in this direction.

[2] *Poset Dimension: Various Definitions*, http://www.math.uni-bielefeld.de/∼hemion/dimension.pdf

[3] This question is dealt with more fully in the Appendix.

physicists normally associate with the concept of "space". A poset exists by itself in its own discrete space.

3.2 Distance

There are various possible methods of defining the idea of distance in a poset. For example we could define the distance from a to b to be the number

$$d(a,b) = \#((a_\downarrow \setminus b_\Downarrow) \cup (b_\downarrow \setminus a_\Downarrow)),$$

where the symbol "#" denotes the number of elements in the set. This is the *symmetric difference metric*[4].

But it is perhaps more sensible to concentrate on the set of positions in a poset $X \in \mathfrak{W}$, rather than just the elements themselves. Thinking about the physical world, we can imagine material objects as being large collections of chains of elements, representing the particles which the objects are made of in space and time. The Earth would be a very large collection of such discrete particles. Then we have the whole Solar System, the Galaxy, and even the billions of further galaxies in the observable universe. These elements are certainly not uniformly distributed throughout the space and time which we see. The material is concentrated in very compact regions with extremely dense collections of element, but between these regions there are almost incomprehensibly vast distances which are practically devoid of any elements. Applying our symmetric difference metric would give absurd distortions in comparison with any sensible geometry for describing the physical world.

On the other hand, thinking about the entire observable universe as being a vast, discrete poset, with the partial order given in the usual way: $a < b$ if b lies in the light-cone above a, then it is obvious that the set of positions would form an extremely, almost unimaginably dense and uniform network throughout space and time. Even in the most empty regions in the vast cavities where there are no galaxies for millions of light years in any direction, the density of positions would be nearly as great as in the middle of a star. For these positions would be mainly determined by the far distant galactic material. Only a few of the nearby elements in the star would be associated with a given position.

And so a concept of space-time could be developed using the set of positions as substitutes for the idea of points in the normal space-time continuum of physics. For example, given $a < b$, we might consider all possible chains of positions connecting a to b. Let

$$a = p_1 < \cdots < p_n = b,$$

with p_i being a position in X for all i, be a maximal such chain in the sense that there is no chain containing more elements. Then we can consider $l = \{p_1, \ldots, p_n\}$ to be, in some sense, a straight line in space-time from a to b. Its length is n. The system of lengths which would result from the definition would correspond roughly with the length given by the usual Lorentz metric of relativity theory, and so a close analogue of the usual theory develops. But of course such concepts as "singularities" – "black holes" and what have you – which arise from differential equations, are meaningless in a discrete theory.

We will use this framework, where positions correspond roughly with points in space-time, to think about the theory of gravity in the next chapter.

[4]Note that it is, at first, only a pseudometric. However we have assumed that the posets we are considering are confluent below and extensional. Thus $d(a,b) = 0$ means $a_\downarrow = b_\downarrow$, and so $a = b$.

But questions remain. The idea of first taking the usual continuous space-time, then using some sort of process to change it into something discrete, seems unsatisfactory. It is an artificial, ad-hoc construction. We are no longer thinking about some sort of *ultimately simple* theory of physics. And so in Chapter 6, and in the Appendix, some ideas will be discussed in an attempt to make this framework appear to be more natural.

Chapter 4

Gravity

4.1 Schwarzschild-like spaces

The Schwarzschild metric is used to describe the gravitational field of a stationary mass m (m being proportional to the amount of the mass) concentrated at the center of four-dimensional Euclidean space $\mathbb{R}^4$ (but existing along the axis used to denote time). Using the isotropic coordinate system, the metric is given by

$$\begin{aligned} ds^2 &= \left(\frac{1-\frac{m}{2r}}{1+\frac{m}{2r}}\right)^2 dt^2 - \left(1+\frac{m}{2r}\right)^4 (dr^2 + r^2(d\theta^2 + \sin^2\theta\varphi^2)) \\ &= A(m,r)dt^2 - B(m,r)(dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)) \\ &= A(m,r)dt^2 - B(m,r)(dx^2 + dy^2 + dz^2) \end{aligned} \tag{4.1}$$

Here r is the distance from the time coordinate axis:

$$r = \sqrt{x^2 + y^2 + z^2}.$$

We can imagine that this formula gives us a way of comparing Schwarzschild space with flat Euclidean space. For a region near to a point (x, y, z, t), it is as if time slows down and spacial distances expand in comparison with the original Euclidean measure. There are three spatial directions and one of time so that the volume of space-time is changed by the factor

$$\left(A(m,r) \times B(m,r)^3\right)^{\frac{1}{2}} = \left(1-\frac{m}{2r}\right)\left(1+\frac{m}{2r}\right)^5 \tag{4.2}$$

Put another way, in comparison with the original Euclidean space, more volume has been packed into a given region. Or, in some sense, the "density" of space-time has increased. Thinking about this in terms of the positions of $X \in \mathfrak{W}$ representing points of space-time, we would have an increased density of positions near to the mass.

Given that the mass doesn't change, then m becomes a constant, and so let us simply write $A(r)$ and $B(r)$ to express the variations with the distance from the time axis. Generalizing things, let us say that for any choice of the functions A and B, equation 4.1 gives a *Schwarzschild-like* space.

The local speed of light remains constant, but for a distant observer the speed of light will *appear* to be the same everywhere if we have $A(r) = B^{-1}(r)$, giving Schwarzschild-like spaces of the form

$$ds^2 = B(r)^{-1}dt^2 - B(r)(dx^2 + dy^2 + dz^2) \tag{4.3}$$

Of course the rule represented by formula 4.3 does not give a solution to the equations of the theory of general relativity, except in the trivial case that B is a constant function. On the other hand, some gravitational spaces with metrics of the form 4.3 are indeed sufficiently similar to Schwarzschild space to lead to the same classical predictions as given by the general theory of relativity. In particular, if we take $B(r) = e^{\kappa m/r}$, where κ is an appropriate constant, then (away from some neighborhood of the time axis) we obtain a theory of gravity which satisfies the classical tests of general relativity (advance of the perihelion of Mercury, deflection of light by the Sun, and the gravitational red-shift of light).[1]

4.2 Adding in elements to a given poset

Theorem 4.1. *Let $X \in \mathfrak{W}$ and let the subset $Y \subset X$ with $Y \in \mathfrak{W}$ be such that Y is 'upwardly dense' in X in the sense that for any two elements $a \| b$ of X, there exist elements of Y in $a_{\Uparrow} \setminus b_{\uparrow}$. Assume that $\mathcal{P}$ is a position in X and $\mathcal{P} \cap Y$ is also a position in Y. Then if an element $y \in Y$ is associated with the position $\mathcal{P}$ in X, it follows that y is also associated with the position $\mathcal{P} \cap Y$ in Y.*

Proof. Since y is associated with $\mathcal{P}$, we must have some $u > \mathcal{P}_{\Downarrow} \setminus \{y\}$ with $u \not> y$. But then $u_{\Uparrow} \setminus y_{\uparrow} \neq \emptyset$. □

What is the meaning of this theorem? Let $Y \subset X$ as in the theorem, and let $\mathcal{Q}$ be a position in Y. Then, considered in X, $\mathcal{Q}$ can be expanded to reach a position $\mathcal{P}$ in X. The position $\mathcal{P}$ is defined as follows. The upper cone of $\mathcal{P}$ is the set of all elements of X which are greater than all elements in the lower cone of $\mathcal{Q}$. Then the lower cone of $\mathcal{P}$ is the set of all elements of X which are less than all elements in the upper cone of $\mathcal{P}$. The position $\mathcal{P}$ in X is thus uniquely defined, and we have $\mathcal{P} \cap Y = \mathcal{Q}$. Furthermore, no elements of $X \setminus Y$ can be associated with $\mathcal{P}$. On the other hand, there may be many other positions $\mathcal{P}^*$ in X, with $\mathcal{P}^* \cap Y = \mathcal{Q}$. But each of them must be associated with at least one element of $X \setminus Y$. To summarize then, let $\mathcal{P}(X)$ and $\mathcal{P}(Y)$ be the set of positions in X and Y, respectively. Then there is a unique embedding $\mathcal{P}(Y) \subset \mathcal{P}(X)$, such that the relationship of association between elements and positions is preserved in the subset.

Now let $x \in X$ be given. Consider the set $x^{\mathcal{P}}$ of positions in X with which x is associated. All positions in $x^{\mathcal{P}}$ must be above x. Considering these positions as being embedded in $\mathbb{R}^4$, we may associate them with the light cones above and below various points of space-time. Let us say that the real function $\rho : R_+ \to R_+$ gives the expected density of the positions of X with which x is associated. That is, let $\Delta \subset \mathbb{R}^4$ be some small region along the light-cone from x at a Euclidean distance r from x. Then the expected number of elements of $x^{\mathcal{P}}$ in Δ is approximately given by $\rho(r) \times S(\Delta)$, where $S(\Delta)$ is the volume of the future light-cone from x contained within Δ. What can we deduce about the form of this function ρ?

Choose an upwardly dense subset $Y \subset X$, with $x \in Y$. We assume that the average density of Y in horizontal hyper-planes of $\mathbb{R}^4$ (with respect to the time axis) is reduced by some constant factor ν, when compared with the original set X. Let $\rho_\nu : R_+ \to R_+$ be the density function for the set of positions in Y with which x is associated. But then it is reasonable to expect that the form of the density function ρ_ν is similar to the original density function ρ. That is, we will assume that $\rho_\nu(r) = \rho(\nu \times r)$, for all distances r from x.

[1] The idea is discussed in §18.2 of H. Yilmaz, 'Introduction to the Theory of Relativity and the Principles of Modern Physics', Blaisdell Pub. Co. 1965, and in Yilmaz, *Phys. Rev.*, **111**, (1958), 1417.

If the density of the elements of X in $\mathbb{R}^4$ is reduced by the factor ν, then the similarity of Y to X means that the density of the positions of X in $\mathbb{R}^4$ must be reduced by the same factor. This gives the equation $\rho(r) = \nu \times \rho(\nu \times r)$. Therefore, if the equation holds true for arbitrary choices of ν, we can choose $\nu = 1/r$. This gives

$$\rho(r) = \frac{1}{r} \times \rho(1) = \frac{constant}{r}$$

4.3 Gravitational effects

Let us imagine some region of space-time which is free of any particular concentrations of matter. Within such a region, space-time is – more or less – flat. This corresponds to the idea that the positions of X are more or less homogeneously distributed within that region of $\mathbb{R}^4$. But then we imagine placing some great mass of material into the region: for example the Earth, or the Sun. This mass is composed of very many particles. By introducing the particles, we introduce many new elements into X, and thus many new positions are introduced as well.

Let us say that this additional mass consists of n particles. We first imagine X without these n particles, then they are added back into X, one by one. Without these particles, we have flat space. That is, the positions of X in $\mathbb{R}^4$ are homogeneously distributed, say with a density of d positions per unit volume of $\mathbb{R}^4$.

After adding in the first particle, we have that the density of positions away from the origin, in space-like hyper-planes, becomes approximately

$$d\left(1 + \frac{k}{r}\right),$$

for some constant k. Adding in the second particle, the density of positions is *incrementally* increased by the factor $(1 + k/r)$, giving a total density of

$$d\left(1 + \frac{k}{r}\right)\left(1 + \frac{k}{r}\right) = d\left(1 + \frac{k}{r}\right)^2.$$

Continuing in this manner, we obtain that the density of positions in the full set X is approximately given by

$$d\left(1 + \frac{k}{r}\right)^n \approx de^{k \cdot n/r}$$

But there are n particles, thus the physical mass of the gravitating object is proportional to n, so that we can write

$$de^{k \cdot n/r} = de^{\kappa m/r},$$

for some constant κ, where m is the mass of the object. And thus we have established a correspondence with the ideas in section 4.1.

What about such effects as Lense-Thirring, or gravitational waves? Such things can be calculated using the differential equations of the usual theory. But the discrete theory offers no practical tools for making such calculations and it is difficult to imagine how a calculus of discrete posets could be formulated. On the other hand, the idea of frame-dragging would seem to be a natural phenomenon in the discrete theory, since the idea of space (given by the set of positions) is directly determined by physical matter (given by the elements of the poset).

Chapter 5

Probability

When discussing probabilities in Chapter 1 we imagined a certain 4-dimensional volume of space and time – a box whose space-like and time dimensions were finite. An experiment with various possible outcomes was placed within this box, and outside the box was the rest of the universe. But the overriding idea was that all possible universes, containing both everything within and outside the box, are equally likely. How can this idea be realized in our framework of discrete posets?

Instead of a finite 4-dimensional box in a space-time continuum, an analogous idea in the framework of discrete posets would be to take a related pair of elements $a < b$ and then to consider the space between, namely $a \between b$, the set of elements greater than a and less than b, to represent the interior of the box. The set of elements not between a and b would represent the outside universe. By placing different configurations of elements into the region between a and b we could represent an experiment having various different outcomes. Thus an experiment with a given outcome would be nothing more than a finite collection of elements placed between a and b, together with ordering relations amongst those elements and with other elements in the rest of the poset. As in Chapter 1, we might imagine that the set of all possible such collections of elements could be divided up into different equivalence classes, representing the different possible results of the experiment. By counting how many members each equivalence class contains we would obtain the relative probabilities.

In the discussion of the two slit experiment we imagined a single fixed universe outside the box and then counted how many configurations within the box are compatible with this outside universe. And then we repeated this for all possible outside universes. Thus, in the end, the probabilities were found by counting the total number of universes in each case.

Translating this into the theory of discrete posets raises certain problems. The sets we are considering are infinite so that simple counting makes no sense. Also the idea of a configuration within a box — in this case in $a \between b$ — being compatible with the rest of the poset outside of $a \between b$ makes no sense. What does "compatible" mean here? Therefore we will have to reduce things to finite sets and also we will have to consider the total number of those sets for each configuration, irrespective of any idea of compatibility.

Being more specific, let us consider a few simple examples. Imagine that we have a finite poset containing the elements $a < b$ and with a further n elements which are not in $a \between b$. The first configuration we consider will be that of two unrelated elements x_1 and x_2 in $a \between b$. For each u with $u \| b$ and $u > a$ we have four possibilities for its relationships with x_1 and x_2.

1. $u > x_1$ and $u > x_2$
2. $u > x_1$ and $u \| x_2$
3. $u > x_2$ and $u \| x_1$
4. $u \| x_1$ and $u \| x_2$

There are analogously four possibilities for each v with $v < b$ and $v \| a$. Let us call this configuration C_1.

Another configuration would be three unrelated elements: y_1, y_2 and y_3 Now for each u with $u \| b$ and $u > a$ we have 8 possibilities, representing the condition that u is either above, or not above y_i for each i. This is configuration C_2.

Since there are more possible "universes" containing C_2 than is the case with C_1, are we justified in concluding that C_2 is more probable than C_1? Of course not. This is a case of the simple observation that in general the more elements there are, the more possible combinations. C_1 and C_2 cannot be sensibly compared with one another. What restrictions are necessary in order to allow sensible comparisons?

One idea would be to say that only configurations in $a \between b$ which have the same number of elements should be compared with one another. But this ignores the basic partially ordered structure of the sets we are dealing with. For example consider the configuration C_3 in $a \between b$ which consists of three elements: z_1, z_2 and z_3 with

$$z_1 < z_2 < z_3.$$

For this configuration an element u with $u \| b$ and $u > a$ has just four possibilities for its relationships with the z_i, namely

1. $u > z_3$
2. $u \| z_3$ and $u > z_2$
3. $u \| z_2$ and $u > z_1$
4. $u \| z_1$

This is the same as for the configuration C_1 despite the fact that there are more elements in C_3 than there are in C_1. But as far as elements not in $a \between b$ are concerned, they are similar.

And so as a general principle we will say that two local configurations in $a \between b$ can be compared only if within the larger set with n elements not in $a \between b$ there are the same number of possible different posets.

Given a set of elements in $a \between b$ we will say that the *transitive power set* $T(a \between b)$ consists of all subsets S of $a \between b$ such that given any $x \in S$ we have that for all $y \in a \between b$ with $y < x$ we also have $y \in S$. Therefore given any element S of the transitive power set $T(a \between b)$, an element u with $u \| b$ and $u > a$ could[1] be assigned to be greater than all the elements of S and not related to the elements not in S in $a \between b$. Each of the different elements of $T(a \between b)$ therefore represent different possible relations of u to the elements in $a \between b$. This leads to the general rule:

Given $a < b$ and a number m, then all possible configurations of elements between a and b having precisely m elements in their transitive power sets will be taken to be equally likely.

[1]We are ignoring here possible relations between u and elements below $a \between b$

This rule ignores the size of the set of elements not in $a \between b$ and it also ignores the restrictions which might follow from possible ordering relations between those elements. Also it ignores possible relationships with elements v not in $a \between b$ with $v < b$ and $v \| a$. In fact it completely ignores everything outside of $a \between b$. Nevertheless we will take this rule as justified owing to the fact that the sets in $\mathfrak{W}$ which we are considering are infinite and upwardly separating.

The rule can also be formulated more simply using the language of partially ordered sets:

Given $a < b$ and a number m, then all possible configurations of elements between a and b having precisely m antichains will be taken to be equally likely.

For if u is an element with $u \| b$ and $u > a$ then given any antichain in $a \between b$, we will obtain a possible variation of the poset by taking u to be greater than all elements of the antichain (and thus necessarily also all elements which are less than at least one element of the antichain) and not greater than all other elements of $a \between b$.

Another way of thinking about this is to take the set of all positions of the poset and imagine that this is what is meant by the idea of "space" in physics. If we have a large volume of space, containing many positions, and compare it with a small volume of space then obviously there are many more possibilities for varying the large volume - trying out alternative configurations within it - in comparison with the small volume. But it would be absurd to say that the large volume is thus more probable than the small volume. It only makes sense to compare different variations of a region with one another if they leave the volume of space unchanged.

Given $a < b$, we assume that all of the elements of the poset in $a \between b$ are essential elements. Let V be the set of positions between a and b which are not associated with any of the elements in $a \between b$. That is to say they are positions determined by elements in $b_{\downarrow} \setminus a_{\downarrow}$ which are not greater than a. The positions in V remain given, regardless of how we place possible variations of elements into the space between a and b. Thus we can ignore them when the different possible variations are compared with one another. On the other hand let W be the set of positions determined just by the elements between a and b in a given variation. Our rule is then that if a different variation produces the corresponding set W', then that variation can only be compared with the original one if W and W' contain the same number of positions.[2] Of course this ignores all the positions which are determined by elements in $b_{\downarrow} \setminus a_{\downarrow}$, as well as possible varied elements in $a \between b$. Our assumption is that on average, given that we are assuming no specific knowledge of the details of the elements outside of $a \between b$, then we have no reason to assume that these positions might favor one or another variation, and therefore they can be ignored.

[2]But note that transitive power sets might exist which do not represent the lower sets of positions.

Chapter 6

Physics Without Laws

6.1 An ultimately simple theory?

This chapter contains much speculation. Perhaps some of the ideas might be amenable to mathematical proof. Or perhaps they might be supported by some sorts of computer simulations. And our basic hypothesis might seem, at first, to be completely absurd. Yet I think that, on second thoughts, most of the obvious objections can be dealt with.

Hypothesis. *Choosing* any *random poset* $X \in \mathfrak{W}$ *– with no further conditions at all – will almost certainly produce an accurate model of the physical world.*

How could this possibly be true? Thinking about examples of finite posets shows that all sorts of chaos is possible. Elements are scattered here and there – indeed, *randomly* – giving no particular pattern at all. How can this be reconciled with the fact that the behavior of physical objects obey very definite laws, down to the finest limits of our ability to measure them? Think of the vast, majestic motions of the planets in the solar system. The randomness hypothesis seems absurd.

But let us imagine a coin tossing experiment, where a coin is tossed *infinitely* many times. One possible outcome is that each and every time, heads comes up. More generally, there are an infinite number of possible outcomes where only *finitely* many tails comes up. Yet, according to the theory of probability, if we have a fair coin then the probability that such an unusual thing happens is precisely zero. In fact, the probability that in the limit, the ratio of heads to tails is not precisely 1 is again zero!

In a similar way, we would expect there to be infinitely many posets in $\mathfrak{W}$ which do not look at all like the real world. Yet it might be that the probability is zero that such a poset might be randomly chosen from $\mathfrak{W}$. Or put another way, from the point of view of probability theory, we would like to show that it is almost certainly true (that is, the probability is 1), that a randomly chosen poset in $\mathfrak{W}$ appears to obey the usual laws of physics.

When calculating the probabilities for the coin tossing experiment, the procedure is to only count finite numbers of throws — say the first n — and then extrapolate the probabilities for an infinite number based on these finite samples. Surely an analogous procedure would be called for when examining the probabilities in $\mathfrak{W}$.

Given $X \in \mathfrak{W}$, the obvious idea would be to take sets of the form $a_\uparrow \cap b_\downarrow$, for pairs of elements $a < b$. But then some of the elements of X might be non-essential. Thus we would be justified in saying that we have the same poset X, whether such

non-essential elements are included or not. For this reason, it seems more sensible to consider the completion of the poset, and thus given $a < b$ in X, a measure of 'space' between a and b could be the number of *positions* of X between them. Following the ideas of Chapter 5, we might consider all possible variations of X between a and b which leave the number of positions between a and b constant, and of course such that the varied poset X' is still a member of $\mathfrak{W}$. One might think of this as being the requirement that the amount of *physical space* between a and b should remain constant for all allowed variations.

In the Appendix, the idea of concentrating on *local probabilities* will be developed further, but framed in terms of finite posets – a restriction which can be easily removed. It is shown there that generalized chains of elements are probable and that isolated elements by themselves are improbable, and also, at least on a local level, 4-dimensionality is probable.

6.2 Are chains more probable?

Let $a < b$ in a poset $X \in \mathfrak{W}$ be such that we are comparing various possible configurations of essential elements a and b such that they all have some given number m of positions in $a \between b$. For example let two such collections of elements be $\{v_1 \ldots, v_n\}$ and $\{v'_1 \ldots, v'_{n'}\}$. As described in Chapter 5, if the configuration $\{v_1 \ldots, v_n\}$ has fewer ordering relations amongst its members than $\{v'_1 \ldots, v'_{n'}\}$ then it must follow that $\{v'_1 \ldots, v'_{n'}\}$ has more elements. That is $n < n'$. These are the *essential* elements. Taking a and b to be farther apart in X and choosing ever larger values of m, we will find that there are many more possibilities for configurations which have many related elements as compared with configurations with few related elements.

Unfortunately though, the possibilities for these variations of X are much more varied than is the case in the coin tossing experiment. And the individual variations of essential elements are not independent of one another. Therefore we must resort to speculating about what might be most probable. Perhaps in the future these speculations will be justified by theoretical arguments, or even computer experiments with appropriate models.

6.3 How the positions of elements are determined by other elements

Let $x \in X$ be some element in a poset $X \in \mathfrak{W}$. Then the pair $(x_{\Downarrow}, x_{\Uparrow})$ is an elementary position in X. But now let us consider the pair $(x_{\downarrow}, x_{\uparrow})$ in $X \setminus \{x\}$. If x is an essential element of X, then $(x_{\downarrow}, x_{\uparrow})$ is not a position in the poset $X \setminus \{x\}$. Still, $x_{\downarrow}$ is the lower set of a position in X, namely $(x_{\downarrow}, V)$, where $x_{\Uparrow} \subset V$ and $V \setminus x_{\Uparrow} \neq \emptyset$. Let us call this the position *directly beneath* the element x.

All elements $v \in V \setminus x_{\Uparrow}$ are such that $x \| v$. On the other hand, since $v > x_{\downarrow}$, if v is a lowest element of $V \setminus x_{\Uparrow}$, a variation of X could be performed, adding in the single new relation $x < v$. According to our previous considerations, this will be probable if v is near to x. Therefore, given that the configuration around x is a probable one, we must conclude that all the elements of $V \setminus x_{\Uparrow}$ are far away from x. This means that locally — near to x — the pair $(x_{\downarrow}, x_{\uparrow})$ does correspond with the position $(x_{\downarrow}, V)$ in $X \setminus \{x\}$.

Concentrating on the situation near to x, let us say that the element $a \in x_{\downarrow}$ is associated with the position directly beneath x. That means that there must be some element $b \not> a$, yet with $b > x_{\downarrow} \setminus \{a\}$. If b is nearer to x than is a, then we can perform

a variation, removing the relation $x > a$ (so that in the varied poset, we have $x \| a$), and adding in the new relation $x < b$. The net result is to have exchanged the close unrelated pair $x \| b$ for the more distant unrelated pair $a \| x$. Thus this variation leads to a more probable poset.

In a similar way, it might be the case that there is an element $c \| x$ which is such that all elements of $x_\uparrow \setminus c_\uparrow$ are further from x than is c. In this case, a variation adding in the new relation $c < x$, and removing the relations of x to all elements of $x_\uparrow \setminus c_\uparrow$, would also result in a more probable poset. In both cases we see that it is probable that (as far as is possible without changing their mutual relationships) the elements near to x are related to x.

We can also consider positions *above* the essential element x. Let (U, V) be a position which is greater than x (so that $x_\Downarrow$ is a proper subset of U), such that there is no other position between (U, V) and x. That is, (U, V) is a position *directly above* x. In contrast to the *single* position which is directly beneath x, there may be more than one position directly above x (assuming of course that they are non-elementary).

What possibilities are there for a non-elementary position (U, V) directly above x? Remembering that all positions are only associated with elements beneath the position, we see that we must have $U = x_\Downarrow \cup \{a_1, \ldots, a_n\}$, for some finite number of elements a_i, $i = 1, \ldots, n$, with $a_i \| x$, and then $V = U_\Uparrow$.

The simplest idea would be to simply choose some single element $a \in X$ with $a \| x$ and $a_\downarrow \subset x_\downarrow$. This would give us $U = x_\Downarrow \cup \{a\}$ and $V \subset x_\uparrow$. In general we can expect to have many such elements as a, and so we would have many different positions directly above x.

But is it probable that there are, in fact, many different positions directly above x? Let us examine a position (U, V) whose lower set is of the form $x_\Downarrow \cup \{a\}$. Thus $U = x_\Downarrow \cup \{a\}$ and all elements of V are above both x and a. However, for all further elements $b \| x$ with $b \neq a$, there must be some element $z \in V$ with $z \not> b$.

Given such a b near to x in the sense that $b_\downarrow \subset x_\downarrow$, let us take a lowest $z \in V$ with $z \not> b$. If z is not far away from x, our argument shows that a variation which introduces the new relation $z > b$ is probable. This would bring with it also the new relations $y > b$, for all the elements $y \in z_\uparrow$. The same could be said for other elements $z' \in V$ with $z' \| z$ and $z' \not> b$. Thus we would have to add in the element b to the lower set of our position directly above x.

So the conclusion we draw is that it is most probable that there are relatively few different positions directly above x, and given such a position (U, V), then the set $U \setminus x_\Downarrow$ contains a relatively large (but of course only finite) number of elements.

All of this can be related to the geometry of Minkowski space. If we take $\mathbb{R}^4$ to be partially ordered using the Lorentz metric, then consider the following pair (U, V) of subsets of $\mathbb{R}^4$. Let $L = \{(0, x, 0, 0) \in \mathbb{R}^4 : x \in [-1, 1]\}$, and take $U = \{p \in \mathbb{R}^4 : \exists l \in L, p \leq l\}$, then $V = U_\Uparrow$. This is a position in $\mathbb{R}^4$, yet it is not "localized", in the sense that it can be identified with the double "light-cone" around a specific point of $\mathbb{R}^4$.

Taking our discrete poset X to be embedded in $\mathbb{R}^4$, we see that there are also many non-localized positions in X — at least with respect to this specific embedding in $\mathbb{R}^4$. On the other hand, our arguments have shown that if X is probable, then none of the essential elements of X will have such a non-localized geometry. They can all be identified with specific points of $\mathbb{R}^4$.

6.4 Generalized chains

If it is more probable that nearby elements are related, rather than being unrelated, then it follows that in a typical poset $X \in \mathfrak{W}$, the elements will tend to form discrete chains, the adjacent elements of which are close together. Let $\mathcal{C}$ be typical chain. Perhaps it is infinitely long, or perhaps it is only finite. Let $x_1 < x_2 < \cdots < x_n$ be some finite segment of adjacent elements along $\mathcal{C}$. Now take some other chain $\mathcal{C}'$, disjoint from $\mathcal{C}$, which is sufficiently long that it contains elements less than x_1 and also elements greater than x_n. Given some particular element x_i of $\mathcal{C}$, then if $\mathcal{C}$ and $\mathcal{C}'$ are far apart, we expect to have many elements of $\mathcal{C}'$ being unrelated to the element x_i. On the other hand, if $\mathcal{C}$ and $\mathcal{C}'$ are close together, then there will be fewer elements of $\mathcal{C}'$ which are unrelated to x_i. Does this mean that it is more probable that $\mathcal{C}$ is close to $\mathcal{C}'$?

In fact, our previous argument cannot be applied to chains. Recall that if (U, V) is some position, then if two given elements a and b are related to one another, we can only have the position being associated with at most one of the elements, *a or b*. On the other hand, if a and b are unrelated, then the position could — in addition — be associated with *both* a and b together. So the conclusion was that a configuration with a being related to b would be more probable.

But now take the two chains $\mathcal{C}$ and $\mathcal{C}'$, and again consider some position (U, V) in X. Assuming that the chains are long enough to contain both elements in U, and also elements not in U, then the position can be associated with at most a single element from each chain — either one element from one of the chains, or two elements, namely one element from the chain $\mathcal{C}$ and another element from the chain $\mathcal{C}'$. This is true regardless of whether or not the two chains are close together; in either case, just a single element of each chain is available to be associated with the position.

On the other hand, an argument can be made that a kind of "generalized" chain might be probable. That is to say, given two distinct chains $\mathcal{C}$ and $\mathcal{C}'$, they might be so close together that each element of each chain is only unrelated to a single element of the other chain. Thus, if $x_1 < \cdots < x_n$ is a segment of $\mathcal{C}$ and $x'_1 < \cdots < x'_n$ the corresponding segment of $\mathcal{C}'$, then we have $x_i \| x'_i$ for each i, yet $x_i < x'_j$ and $x'_i < x_j$ if $i < j$. Let us now imagine that $\mathcal{C}$ is near to $\mathcal{C}'$, in the sense that both $x_{i\downarrow} \setminus x'_{i\downarrow}$ and $x'_{i+1\downarrow} \setminus x_{i\downarrow}$ have few elements, for each i. In this case it is unlikely that a randomly chosen position (U, V) in X will be associated with both an element of $\mathcal{C}$ and also an element of $\mathcal{C}'$. Instead, just a single element from the union of the two chains $\mathcal{C} \cup \mathcal{C}'$ would be more likely. Therefore such a configuration where two chains run very closely parallel with one another would be probable.

More generally, the argument shows that generalized chains of the form $\{\ldots, x_{-1}, x_0, x_1, x_2, \ldots\}$, with the relations generated by $x_i < x_{i+m}$, for all i, and for some fixed $m > 1$, yet $x_i \| x_j$ for $i < j < i + m$, would also be probable. Let us call this a *generalized chain of order m*.

6.5 The geometry of generalized chains

How would such a generalized chain be embedded in $\mathbb{R}^4$? After all, the individual elements of the chain are essential positions of X, and so they should correspond with points of $\mathbb{R}^4$.

Let us consider a generalized chain of order 4. Its elements are

$$\ldots, x_{-1}, x_0, x_1, \ldots$$

with $x_n < x_{n+4}$, for each n. But $x_n \| x_{n+1}$, $x_n \| x_{n+2}$, $x_n \| x_{n+3}$ and $x_n \| x_{n+4}$. In particular if we concentrate on the elements x_1, x_2, x_3 and x_4, we see that they are all unrelated to one another. But the element x_8 is greater than all of them. In fact x_m is greater than x_i, for $i = 1, 2, 3, 4$, if and only if $m \geq 8$. Thus each of the four elements x_1, x_2, x_3 and x_4 must be in the lower set of the position directly beneath x_8. But then, since the justification for having generalized chains was based on the idea that the elements are as close together as possible, we see that it would be reasonable to assume that x_8, considered as a point in $\mathbb{R}^4$, is near to the lowest possible point above the 4 points where x_1, x_2, x_3 and x_4 are situated. Then x_9 would be near to the point of $\mathbb{R}^4$ which is the lowest point above the four unrelated points x_2, x_3, x_4 and x_5. And so forth. Thus we obtain a spiraling chain of points, such that the spiral as a whole proceeds upwards through $\mathbb{R}^4$ in a straight line.

Generalized chains of orders other than 4 are also conceivable. However 4 fits so naturally into 4-dimensional space-time that we are tempted to associate it with the electron lines in Feynman diagrams. Perhaps chains of higher orders could represent quarks, with their "hidden" dimensions.[1]

6.6 Particles and antiparticles

Consider an essential element $x \in X$, which we take to be an element of a generalized chain and let (U, V) be the position directly beneath x. The fact that x is essential means that $V \setminus x_{\Uparrow}$ is not empty. On the other hand, since X is probable, the elements of $V \setminus x_{\Uparrow}$ are far from x. Thus, at least within a reasonably large neighborhood of x, we have $x_{\downarrow} = U$ and $x_{\Uparrow} = V$.

If the generalized chain is of order n then the next $n - 1$ elements of the generalized chain are unrelated to x. Let us say that x is the zeroth element of the generalized chain; call it x_0. Then $x_1, \ldots, x_{n-1}$ are unrelated to x_0. But the elements $x_n, \ldots, x_{2n-1}$ are all greater than x_0, and they are all unrelated to one another. Then going downwards through the generalized chain, we find that the elements $x_{-2n+1}, \ldots, x_{-n}$ are unrelated to one another.

Let $C_- = \{x_{-2n+1}, \ldots, x_{-n}\}$ and $C_+ = \{x_n, \ldots, x_{2n-1}\}$. Therefore the element x_0 lies between C_- and C_+, and we have $C_- \subset U$ and $C_+ \subset V$. But we can imagine that not only the position (U, V) lies between C_- and C_+. There may be numbers of other positions as well. That is to say, there is a certain amount of room for adjusting the exact position of the element x_0 within the generalized chain.

Nevertheless, the elements of the generalized chain should be closely spaced. One way to do this is to require that x_0 be nearly as low as possible in the ordering of X, while still being above C_-. If this is true of x_0, then in order to keep the generalized chain being closely spaced, we would have the next element, x_1, also being as low as possible with respect to the highest elements in the generalized chain beneath x_1, and so on. Call such a generalized chain a "particle".

Alternatively, the close spacing of the generalized chain might be achieved by having x_0 as high as possible, yet beneath C_+. And this requirement would hold for the other elements in the generalized chain. In this case we would have an "antiparticle".

[1] And thus they would not be so nearly associated with localized points in an embedding $X \to \mathbb{R}^4$. That is, for such higher order generalized chains, the embedding would not be exactly order preserving. But after all, we expect Minkowski space to give only an approximation of the exact ordering structure of X.

6.7 Photons

We have seen that an element x_0 of X, situated somewhere along a generalized chain, if it is a particle, can be thought of as being as low as possible with respect to C_-, the set of highest elements of the generalized chain beneath x_0. This would certainly be the case if the position directly beneath x_0 was associated with the elements of C_-, and no others. Given that this is also true of the elements along the generalized chain near to x_0, then as we have seen in the last section, the generalized chain would follow a straight path through $\mathbb{R}^4$. But what would then happen if the position directly beneath x_0 was associated with all but one of the elements of C_-, and to compensate for this, the position is associated with some other element $y \in X$, which lies on some other generalized chain?

Let us say that it is the element $c \in C_-$ which — unexpectedly — is not associated with the position directly beneath x_0. Since the generalized chain should be as closely spaced as possible, we would expect that the next element in the generalized chain, namely x_1, would be such that the position directly beneath x_1 would be associated with c. Then, continuing the pattern of the generalized chain on from x_1, we will find that the sequence of the spiral in the generalized chain will have changed. One can think of this as being a change of "spin".

What is the effect of all this on the generalized chain which contains y? Let D_- denote the set of highest elements of this generalized chain beneath y, and D_+ is the set of lowest elements above y. If y is on a generalized chain representing an antiparticle, then it is as high as possible, while being constrained to be beneath D_+. Yet the relationship with x_0 gives a further definite constraint. So we would expect the interaction to be such that now y is the highest element beneath all but one of the elements of D_+, and instead of this one element left out of D_+, we have x_0. Therefore the generalized chain containing y also experiences a change of spin.

An interaction is also possible if the y generalized chain represents a particle, or if the x generalized chain represents an antiparticle. For the elements x_0 and y_0 can be determined by a combination of elements both above, and below. They are not subject to the constraints we have when the determining elements are exclusively within the single generalized chain.

This relationship between pairs of elements of different generalized chains would explain the idea of a photon — or a gluon, and so forth in the case of generalized chains of higher order. Within this picture, a photon is not a particle, in the sense of it being represented by some sort of chain of elements in X. Nevertheless, such a photon would transfer spin, as expected.

But more than this; photons (also gluons, and whatever similar higher order "exchange" particles there might be) exert forces on the true, massive particles involved in the photon exchange. That is to say the directions of the particles are altered after such a particle exchange. Thinking about Feynman diagrams, this is just what we see. The points where the photons are emitted or absorbed are vertices of the diagrams. Between these vertices, the electron or positron edges in a diagram are straight lines.

We still haven't accounted for the fact that pairs of electrons repel one another, while electrons and positrons attract one another. But an argument can again be made to explain this within our model. Namely the exchange of a photon will allow the corresponding generalized chains to be more closely spaced at the point where the photon is emitted or absorbed, owing to the fact that an element of the generalized chain is left out of the sequence. If the photon connects a particle with an antiparticle then this effect will be greater than if the photon connects two like particles. However, in either case a closer spacing will result. Why is it then that the weaker effect leads

to the particles repelling one another? This will follow from the fact that on the whole, the density of X remains unchanged. Thus a stronger attraction will result in a true attraction within the geometry of X, while a weaker attraction will appear as a repulsion.

6.8 The density of discrete space

In Chapter 4 we considered a poset embedded in an order preserving way in 4-dimensional Minkowski space. The set of positions of the poset could also be represented by points in Minkowski space, preserving their ordering as well. We assumed that these positions, mainly determined by the far distant matter in the universe, were distributed more or less uniformly in space-like hyperplanes perpendicular to the time axis. Furthermore in a region near to the origin, the density of the positions — that is the average number of positions in 4-dimensional cubes of the Euclidean space — was assumed to be approximately constant.[2] Then adding in an extra collection of elements representing a gravitational body situated along the time axis near the origin, we deduced that extra positions would be added into the poset, and thus into Minkowski space, changing the density of the points representing these positions in the region by something of the form

$$de^{\kappa m/r}$$

And so, considered with respect to 4-dimensional cubes in Euclidean space, the density of the positions near the origin has increased.

Another way of looking at this, as an alternative to thinking about 4-dimensional cubes, would be to consider say two points p and q in Minkowski space with $p < q$, say a distance -1 apart when taken with respect to the Lorentz metric. Or, to avoid being too pedantic, let us just say that they are a unit distance apart. Then we consider the number of positions between p and q to be a measure of the density of positions near to those points.

But as we have seen, according to the theory of general relativity, in the presence of the gravitational body the metric is altered, and consequently the ordering relations in the poset are altered as well, particularly those near to the gravitational body. For p and q to be a unit distance apart in this new, altered metric, they must be stretched further apart in the underlying Euclidean space, to the points p' and q'. That is to say, the Euclidean distance from p' to q' is greater than that between p and q. On the other hand, in this altered metric the space between p' and q' — that is the intersection of the light-cone above p' with the light-cone below q' — has become narrower so that all in all, the volume of Euclidean space between p' and q' is less than the volume between p and q. In this smaller volume there are fewer positions. Thus if we choose the gravitational constant appropriately we will find that when the gravitational mass has been added in, the number of positions between p' and q' using the metric of general relativity will be nearly equal to the number of positions between p and q in the original poset with the Lorentz metric. Or put simply, the density of positions remains approximately constant, regardless of whether the gravitational mass is included or not.

Up to now we have been thinking about the idea of the density of positions by relating this to some assumed embedding in 4-dimensional Minkowski space. But

[2] Of course, keeping in mind our definition of strong discreteness, if we were to think of embedding the whole of the poset in Minkowski space then the density of positions would have to decrease strongly when going very far backwards along the time axis.

there is a simple way to define this density purely in terms of the poset itself. Let $\mathcal{P}$ and $\mathcal{Q}$ be two positions with $\mathcal{P} < \mathcal{Q}$. Take n to be the number of elements of the poset in the set difference $\mathcal{Q}_{\Downarrow} \setminus \mathcal{P}_{\Downarrow}$ and let N be the number of positions between $\mathcal{P}$ and $\mathcal{Q}$. Then we define the density of the poset between $\mathcal{P}$ and $\mathcal{Q}$ to be the ratio N/n.

This definition depends on the particular positions $\mathcal{P}$ and $\mathcal{Q}$ which were chosen and is thus overly special to just those two positions. In order to be able to compare the density in one region of the poset with another we might fix the number n and then compare the density in one region with another region using some different pair of positions, $\mathcal{R} < \mathcal{S}$, with precisely the same number n of elements in $\mathcal{S}_{\Downarrow} \setminus \mathcal{R}_{\Downarrow}$. We obtain some number N' of positions between $\mathcal{R}$ and $\mathcal{S}$. Then we compare the ratios N/n and N'/n.

Of course this method does not give a definite, unambiguous number which we could call the absolute density of positions at some given position in the poset. It depends on the number n we choose and the particular positions chosen for counting what is between them. Nevertheless, assuming that n is not too large so that it reflects the idea of looking at local densities, we might consider some sort of average of all possible such densities at a given position, where the average is taken over all pairs of positions with respect to the number n with the given position between them.

The assertion is then that for randomly chosen posets in our class $\mathfrak{W}$, it is almost always the case that the density is nearly constant throughout the poset, and this would give an explanation of gravity analogous to that of general relativity.

Indeed, this idea of a nearly constant density of positions is only natural. After all, most of the elements associated with a given position are in the far distant universe. And given some set of n elements around the periphery of some $\mathcal{Q}_{\Downarrow} \setminus \mathcal{P}_{\Downarrow}$, we would expect the different possible combinations of those elements to be associated with nearly the same number of positions, regardless of where in the poset the positions $\mathcal{P}$ and $\mathcal{Q}$ were taken. Anything else would result from some very improbable special configuration of those very distant unrelated elements.

Chapter 7

Is the Discrete Model Reasonable?

It is natural to consider posets as a whole, rather than extracting small subsets and ignoring the rest. In physics, the corresponding idea would be the subject of cosmology.

The standard view of cosmology is that everything started with a big bang about 15 billion years ago. Since then everything has been expanding, and we can see that this expansion continues since the light from distant galaxies appears to us to be red-shifted. This is attributed to a Doppler-shift effect. Since matter and energy, taken together, can neither be created nor destroyed, it follows that the universe is becoming less and less dense as time progresses. Despite this, the universe is filled with black holes of various sizes. In particular, galaxies form around large black holes, sited at their cores. Perhaps there is also a dense assemblage of small black holes, making up the mysterious "dark" energy and matter, which is needed in order to explain the rotational properties of galaxies — under the assumption that gravity is the main force for determining the shape of a galaxy.

The discrete model we have been describing fails to have almost all of these things.

- In order to reduce the number of defining properties, or axioms, to a minimum, it was found to be sensible to assume that the underlying poset has no extreme elements. In particular there is no single minimal element (that is, no "big bang" element).

- The red-shift in the model arises from the asymmetry which results from having $x_\downarrow \setminus y_\downarrow$ being finite, while $x_\uparrow \setminus y_\uparrow$ is infinite, for any two elements $x \| y$. Thus the measure of time changes if we progress upwards in the ordering of the poset. The difference in the measure of time would result in the illusion of a red-shift of light from distant galaxies.

- The model does not include an abstract law of conservation of mass-energy. On the contrary, these things can only be measured with respect to an imaginary background "space-time continuum" In the discrete model, the measure of space-time is determined by the matter contained in it — the elements of our poset. Thus the density of this mass-energy is determined by the mass-energy itself, hence it remains more or less constant throughout the poset.

- The idea of black holes is untenable in any discrete model.

- Thus there would be no apparent limit to the mass of a neutron star. The concept of super-massive black holes would be replaced by super-massive collapsed stars. Ancient, inert collapsed stars might exist in large numbers, emitting little visible light.

- Finally, an argument can be made that in a region of the poset which has few elements, the positions directly beneath essential elements are not as constrained to be "round" in comparison with essential elements which are surrounded by a dense configuration of essential elements. That is, following the argument of the last chapter, they may depart somewhat from having a double cone structure in an embedding in $\mathbb{R}^4$. Thus in a relatively empty region of space, more positions come into consideration for determining individual essential elements in comparison with regions containing many essential elements. This could produce the illusion of "dark matter".

Although the discrete model departs from the standard differential manifold model in all these points, still, the observational evidence provides no basis for ruling either of these models out. However one possible, practical test does suggest itself. This test would involve the observation of pulsars. The newest observations seem to have established beyond question the idea that the pulsing of pulsars is due to rotation, not to any sort of radial pulsation. Given that, then the general theory of relativity places an upper limit on the possible periods of pulsars. The millisecond pulsars are very near to this limit already. Therefore, if a pulsar with a still much shorter period of pulsation were to be found, it would falsify general relativity

Our arguments have been motivated by Feynman's formulation of quantum electrodynamics in terms of his Feynman diagrams. Thus, if the arguments could be shown to have substance, then the discrete theory would be nothing more than a justification for the mathematical methods used when working with such diagrams. Perhaps the quarks can be shown to arise as an expression of generalized chains of order greater than 4 in the discrete model.

Appendix A

Local Probabilities in Finite Partially Ordered Sets

A.1 Definitions

A set a is *transitive* if for any element $b \in a$ which is itself a set containing an element $c \in b$ we have also $c \in a$. We will consider transitive sets which are such that all elements which are sets are themselves transitive[1]. Call this hereditarily transitive. In other words, this is a way of describing partially ordered sets such that the ordering becomes set inclusion. Thus we might interchangeably use the notation $a < b$ or $a \in b$ to mean the same thing. An element a in a partially ordered set can then be thought of alternatively as the set of all elements which are less than or equal to a, allowing us sometimes to simplify the notation.

Let X be a partially ordered set. For elements $a, b \in X$ we use the notation $a \| b$ to mean that both $a \not\leq b$ and $b \not\leq a$, and $a \perp b$ to mean that either $a \leq b$ or $b \leq a$. Given $a < b$ then $a \between b$ is the set of elements $\{c \in X : a < c < b\}$.

As a first idea of what we mean by local probabilities, consider the following. Let n and N be two numbers with $n < N$. We then take all possible partially ordered sets X consisting of N elements, and for each such X we consider all possible pairs $x < y$ of elements in X such that $x \between y$ consists of precisely n elements. These different sets of the form $x \between y$ fall into equivalence classes, where two such sets are equivalent if they are isomorphic as partially ordered sets. Then we can say that the relative probabilities of the different equivalence classes of these "between sets" is proportional to the number of sets in each class. Given n, the relative probabilities will depend upon the choice of N. Perhaps in the limit as N becomes larger, these relative probabilities might converge to some limiting values.

But there is a problem with this method. To illustrate the problem, consider the situation with $n = 2$. That is, we have two elements $x < y$ and two further elements u and v with $x < u < y$ and $x < v < y$. There are two cases: either $u \| v$ or $u \perp v$. In the case $u \perp v$, we either have $u < v$ or else $v < u$. But both are isomorphic as partially ordered sets, so to be definite we choose $u < v$.

Let us take $x \between y$ to be contained within some larger partially ordered set X which also contains both x and y, and which is such that $u < v$ in X. Thinking in terms of hereditarily transitive sets, between x and y we have the two different sets u and v, where of course $u \in v$.

[1] Thus they are identical with their transitive closures.

On the other hand let X' be identical with X except that we now have $u \| v$. This gives *three* different sets between x and y, namely: u, v, and $u \cup v$ in X'. A correspondence between X and X' could be given by the following rule: An element s of X' which is not between x and y (and therefore s is an element of X as well) considered as a set, is to contain the set $u \cup v$ if and only if $v \in s$ in X. Furthermore, $u \in s$ in X' if and only if $u \in s$ in X. On the other hand, s is contained in $u \cup v$ if and only if $s \in v$ in X, and $s \in u$ in X' if and only if $s \in u$ in X.

But this leaves us with the additional set v in X'. Therefore by rearranging the elements of X' we can generally obtain new sets X'', containing the configuration $u \| v$, which do not correspond with any rearrangement of X containing the configuration $u < v$. That is, the configuration $u \| v$ gives us, in effect, three elements to work with, while the configuration $u < v$ has only two.

Obviously there are always more possible combinations of three objects than there are of two. It would be nonsense to thus conclude that three things are always more probable than two. One must compare like with like. In our case we take this to mean that it is only fair to compare two different possible configurations between x and y if they have the same number of elements in their power sets. (Where we assume that an element of the power set, considered as a union of sets in $x \between y$, contains all the elements in each of those sets.) This leads to the definition, which we formulate in terms of hereditarily transitive sets:[2]

Definition 5. *Let n be a given number and let $x \in y$ be two elements. Denote by Ψ the set of all different hereditarily transitive sets between x and y which are such that there are precisely n elements in their power sets. If there are m members of Ψ then we will define each to have the relative probability $1/m$.*

This definition involves a number of assumptions. No mention is made of the larger partially ordered sets within which the possible configurations between the elements x and y are supposed to exist. Our assumption is that the number of possible ways a configuration can be embedded in the total of all the various larger partially ordered sets is the same for all configurations with a given number of elements in their power sets. Or put another way, it is the assumption that the relative probabilities for different configurations, taken with respect to possible partially ordered sets of N elements in which they are contained, converge to a limit as $N \to \infty$.

A.2 Generalized chains

Let some region $x \between y$ be given and let us consider the case that $x \between y$ consists of n unrelated elements. (As always, we will assume that all elements are themselves sets.) That is

$$x \between y = \mathcal{U} = \{u_1, \ldots, u_n\}$$

such that for any $i \neq j$ we have $u_i \| u_j$. Then the power set consists of 2^n subsets.

Another possible configuration for $x \between y$ might be a chain of length $2^n - 1$. That is

$$x \between y = \mathcal{V} = \{v_1, \ldots, v_{2^n - 1}\}$$

[2]Strictly speaking we should consider the number of elements in the MacNeille completion of the poset between x and y. That is, using the notation of the main text, the set of positions between x and y. But the given argument in terms of hereditarily transitive sets is simpler and it suffices in general. And for the application in the main text it is appropriate, considering that we have shown that in $\mathfrak{W}$, the elements associated with any given position always lie below that position.

with $v_i \in v_{i+1}$ for all $1 \leq i \leq 2^n - 2$. Again, there are 2^n subsets in this configuration, so that it has the same probability as does $\mathcal{U}$.

A third possibility for $x \between y$ consists of $2^n - 2$ elements which form a simple *generalized* chain, namely

$$x \between y = \mathcal{W}_1 = \{w_1, \ldots, w_{2^n-2}\}$$

with $w_1 \| w_2$, $w_1 \in w_3$, $w_2 \in w_3$, and then $w_i \in w_{i+1}$ for all $3 \leq i \leq 2^n - 3$. Once again the number of subsets is 2^n so that $\mathcal{W}_1$ has the same probability as the other two configurations which we have considered so far.

But then we also have

$$x \between y = \mathcal{W}_2 = \{w_1, \ldots, w_{2^n-2}\}$$

with $w_2 \| w_3$, $w_1 \in w_2$, $w_1 \in w_3$, $w_2 \in w_4$, $w_3 \in w_4$, and then $w_i \in w_{i+1}$ for all $4 \leq i \leq 2^n - 3$. Again, the number of subsets is 2^n.

In fact, following this pattern one sees that we have $2^n - 3$ such generalized chains $\mathcal{W}_1, \ldots, \mathcal{W}_{2^n-3}$, all of which have 2^n subsets, and thus they are all equally probable.

Many further configurations fit into this pattern. For example we have generalized chains with two pairs of unrelated elements: $w_i \| w_{i+1}$ and $w_j \| w_{j+1}$ for $|i - j| \geq 2$ and the rest of the elements forming a simple chain. If $i = j - 2$ we might also consider $w_i \in w_j$, while $w_{i+1} \notin w_j$, and so on. There are nearly n^2 such configurations. And then we can have generalized chains with three pairs of unrelated elements. There are nearly n^3 such configurations, although we do notice that the number of elements in the configuration is reduced by one for each such addition.

Once we have exhausted all these possibilities we can then consider generalized chains having various numbers of unrelated triples, quadruples, and so forth. Eventually we come back to our single original set $\mathcal{V}$ consisting of n unrelated elements. All of these configurations are equally probable.

If we restrict ourselves to just the types of configurations which have been described here, it is obvious that if the number n is reasonably large and a configuration $\mathcal{C}$ for the set $x \between y$ is chosen at random, then we expect the height of $\mathcal{C}$ to be much greater than its width. Here the height is taken to be the length of the largest possible simple chain, and the width is the number of elements in the largest possible subset consisting of mutually unrelated elements. Therefore $\mathcal{C}$ resembles a chain; a fuzzy, or generalized chain.

A.3 Other structures

The generalized chains considered in the last section are characterized by having a limited width such that for each element u in the chain, the number of elements unrelated to u is limited. If the number of elements we are considering is much greater than these limits then other structures than a single generalized chain might be probable.

As an example let us consider two simple chains, of height k and l:

$$U = \{u_1, \ldots, u_k : u_i \in u_{i+1}, \forall 1 \leq i < k\},$$

$$V = \{v_1, \ldots, v_l : v_i \in v_{i+1}, \forall 1 \leq i < l\}.$$

If $u_i \| v_j$ for all i and j then the number of subsets in the power set is simply $k \times l$. Thus if we are to compare this with a single chain, that chain would have to have

height $k \times l$. But in general there may be some relations of the form $u_i \in v_j$ or $v_i \in u_j$. How many possible different relations and how many subsets can there be?

To begin to estimate how many different possible relations there are, let us consider the possibilities for relations of the form $u_i \in v_j$. The thing to note is that if $u_i \in v_j$ then we must also have $u_{i-1} \in v_j$ and also $u_i \in v_{j+1}$. Let is write $F(k,l)$ to represent the number of possible relations here. Then we see that if $u_k \in v_l$, it is clear that the number of possible relations for the further $l-1$ elements of V is given by $F(k,l-1)$. If $u_k \notin v_l$, but $u_{k-1} \in v_l$ then the number of possible relations for the further $l-1$ elements of V is $F(k-1,l-1)$. Proceeding down through U, we obtain the recursive formula:

$$\begin{aligned} F(k,l) &= \sum_{t_1=1}^{k} F(t_1, l-1) \\ &= \sum_{t_1=1}^{k} \sum_{t_2=1}^{t_1} F(t_2, l-2) \\ &\vdots \\ &= \sum_{t_1=1}^{k} \sum_{t_2=1}^{t_1} \cdots \sum_{t_{l-1}=1}^{t_{l-2}} F(t_{l-1}, 1) \\ &= \sum_{t_1=1}^{k} \sum_{t_2=1}^{t_1} \cdots \sum_{t_{l-1}=1}^{t_{l-2}} t_{l-1} \end{aligned}$$

since $F(t_{l-1}, 1) = t_{l-1}$.

We can bound this sum from below by taking integrals.

$$\begin{aligned} \sum_{t_1=1}^{k} \sum_{t_2=1}^{t_1} \cdots \sum_{t_{l-1}=1}^{t_{l-2}} t_{l-1} &> \int_0^k \int_0^{t_1} \cdots \int_0^{t_{l-2}} s_{t_{l-1}}\, ds_{t_{l-1}} \ldots ds_2 ds_1 \\ &= \int_0^k \int_0^{t_1} \cdots \int_0^{t_{l-3}} \frac{{s_{t_{l-2}}}^2}{2}\, ds_{t_{l-2}} \ldots ds_2 ds_1 \\ &= \int_0^k \int_0^{t_1} \cdots \int_0^{t_{l-4}} \frac{{s_{t_{l-3}}}^3}{3*2}\, ds_{t_{l-3}} \ldots ds_2 ds_1 \\ &\vdots \\ &= \frac{k^{l-1}}{(l-1)!} \end{aligned}$$

Stirling's formula is

$$(l-1)! \approx \sqrt{2\pi(l-1)} \left(\frac{l-1}{e}\right)^{l-1},$$

giving

$$F(k,l) \gtrapprox \frac{1}{\sqrt{2\pi(l-1)}} \left(\frac{ke}{l-1}\right)^{l-1}.$$

Assuming that we have chosen the chain U to be not shorter than V, then we have $k \geq l$ and so $F(k,l)$ grows very rapidly as the height of the chains grows. (Note that even if $l > k$, the number $F(k,l)$ will still be large despite the fact that our integral approximation will now underestimate things drastically.)

In addition to the possible relations of the form $u_i \in v_j$, we also have the relations of the form $v_i \in u_j$ to take into account. For each possible configuration considered in the calculation of $F(k,l)$ we have many such further relations, the number of which should be multiplied by $F(k,l)$ to obtain the total number of different configurations. But note however that we can never have a pair with both $u_i \in v_j$ and at the same time $v_j \in u_i$; a set cannot be contained in one of its elements.

These different configurations generally have different numbers of subsets in their power sets so that they cannot be compared with one another. However if we consider configurations with relatively large numbers n of elements then the equivalence classes of the configurations which are comparable will generally be large. This follows since, as we have seen, there are less than n^2 possible such equivalence classes, yet many more possible configurations.

The conclusion is that if n is reasonably large, and if we restrict ourselves to configurations having either one or two simple chains, then it is overwhelmingly probable that we will have a configuration with two simple chains.

But it is also obvious that there are many more possibilities besides these. For example we might consider generalized chains whose width is limited by some fixed value w. For n large in comparison to w, any element of the generalized chain is not related to only a small number of further elements in the chain. Thus we can think of the generalized chain as consisting of a large number of segments which act somewhat as if they were elements of a simple chain. We can then compare the probability of having a single generalized chain with two interacting, but separate shorter generalized chains. We can apply our argument in this case as well to show that for large n, it is probable that we will have two separate generalized chains. Extending the argument to larger numbers of generalized chains, we see that as n increases, it is probable that we will have proportionally more of these interacting generalized chains.

Are there other structures besides generalized chains, or at least configurations constructed from pieces of generalized chains, which are probable as local structures in finite partially ordered sets? Perhaps not.

A.4 Dimension

As in the last section, let us assume that we have two distinct, simple chains of height p and q:

$$U = \{u_1, \ldots, u_k : u_i \in u_{i+1}, \forall 1 \leq i < p\},$$

$$V = \{v_1, \ldots, v_l : v_i \in v_{i+1}, \forall 1 \leq i < q\}.$$

We would like to examine a region which can be "indexed" by the elements along U and V in the following manner.

Let x be some element in this region. Then there are four unique elements, $\{u_i, u_j, v_k, v_l\}$, two along U and two along V, which index the element x. The two elements u_i and u_j on U indexing x are such that $x \in u_i$ but $x \notin u_{i-1}$ and $u_j \in x$ but $u_{j+1} \notin x$. The elements along V are similarly such that $x \in v_k$ but $x \notin v_{k-1}$ and $v_l \in x$ but $v_{l+1} \notin x$. We assume that the elements in the region are uniquely represented by this indexing. That is, if x' is some other element in the region indexed by the four elements $\{u_{i'}, u_{j'}, v_{k'}, v_{l'}\}$, then at least one of those elements is different from the indexing of x.

We are interested in the question of whether or not x and x' are related to one another. Given the way we have chosen the index elements, we see that if $x \in x'$ then we must have $u_i \subseteq u_{i'}$, $u_j \subseteq u_{j'}$, $v_k \subseteq v_{k'}$, and $v_l \subseteq v_{l'}$. (Here we use the

notation "$a \subseteq b$" to mean the same thing as "$a \in b$ or $a = b$".) Or, in other words, the index elements for x are all *less than or equal to* the respective index elements for x'. For all other relationships of the indexing elements, we must have $x \notin x'$. A similar condition holds for $x' \notin x$. If both conditions hold, then we must have $x \| x'$.

Therefore we must have $x \| x'$ in all cases *except* when the indexing elements of x are all less than (or they are all greater than) or equal to the respective indexing elements of x'. But if, say, the indexing elements of x are all less than or equal to the respective indexing elements of x' then the situation is unclear. We might have either $x \in x'$ or $x \| x'$.

So let us consider the case that all the respective indexing elements of x are less then or equal to those of x'. We then have two different possible configurations: $W_\in$ where $x \in x'$ and $W_\|$ where $x \| x'$. How many subsets does $W_\in$ have in comparison with $W_\|$? The answer is that there are more subsets in $W_\|$ than there are in $W_\in$.

To see this, begin by observing that most of these subsets are the same, both for $W_\in$ and for $W_\|$. Only those which contain x and/or x' might be different. So let us assume that there are m elements along U between $u_{j'}$ and u_i, and furthermore we assume that there are n elements along V between $v_{l'}$ and v_k. (If either m or n were to be zero then we must have $x \in x'$, and so the case $x \| x'$ would not occur.) Each of the subsets of $W_\in$ corresponds with a subset of $W_\|$. In particular each subset in $W_\in$ of the form $u_s \cup v_t \cup x$, for $j' < s < i$ and $l' < t < k$, we have the corresponding subset $u_s \cup v_t \cup x$ in $W_\|$. Similarly, for each subset in $W_\in$ of the form $u_s \cup v_t \cup x'$, we have the corresponding subset $u_s \cup v_t \cup x \cup x'$ in $W_\|$. But then in addition to these, we have the $m \times n$ subsets of the form $u_s \cup v_t \cup x'$ in $W_\|$ (each of which *do not* contain the element x), and there are no corresponding subsets to these in $W_\in$.

Therefore, since $W_\in$ and $W_\|$ contain different numbers of subsets, they cannot be compared with one another. But, as before, it is possible to add in some extra elements to $W_\in$ in various ways, producing an expanded version $W_\in^*$ of $W_\in$ which does have the same number of subsets as does $W_\|$. For example, we could lengthen the chain U in $W_\in$ by attaching $m \times n$ new elements $\{u_{p+1}, \ldots, u_{p+m\times n}\}$ with $u_p \in u_{p+1}$, and then $u_i \in u_{i+1}$, for all i between 1 and $p + m \times n$. To complete the picture, we assume that also $v_q \in u_{p+1}$.

There are many other possibilities for adding new elements into $W_\in$. For example we could add elements to the other chain V, to both, or midway along the chains at various positions, adjusting the number of new elements in each case so that the total number of subsets remains constant. Therefore we conclude that it is very probable that $x \in x'$ if the index elements of x are all less than or equal to the corresponding index elements of x', becoming overwhelmingly probable when the indexing chains are long.

So let us assume that we have two distinct chains, U and V, and also many elements in a region which are indexed by these chains such that for any two of these elements, the indexing is not identical, and furthermore, given two such elements a and b, we have $a \in b$ if and only if the indexing elements for a are all less than or equal to the corresponding indexing elements for b. Then we conclude that the set of these elements, considered as a partially ordered set, is 4-dimensional.

Recall the definition of dimension within the theory of partially ordered sets. Each partial order for a given set can be expanded by adding in further ordering relations to obtain a totally ordered set which contains the original partial order. A *realizer* of the partial ordering is a collection of total orders, each of which contains the original partial order, such that the original partial order is the intersection of all the ordering relations in the realizer. A partially ordered set has the dimension n if there is a realizer consisting of n totally ordered sets, where n is the smallest such number.

Applying this to our situation with the two chains U and V, we can find a realizer consisting of just 4 totally ordered sets. They are obtained by using the ordering of each of the four indexing elements. For example, given that $a \| b$, the first total ordering involves adding in the relation $a < b$ if the upper indexing element along U of a is contained in the upper indexing element of b along U. In this way we obtain the first of our total orderings. The other three are obtained similarly.

When thinking about this argument, it might be objected that the same ideas could be applied to the situation with just one single indexing chain, thus seemingly leading to the conclusion that we would have 2 dimensions rather than 4. That is to say, let the single chain

$$U = \{u_1, \ldots, u_k : u_i \in u_{i+1}, \forall 1 \leq i < p\}$$

be given, together with two elements x and x' not in the chain, but such that the chain has elements which contain both x and x' and also elements contained in both x and x'. As before, we take the indexing elements for x to be u_i and u_j, and for x' to be $u_{i'}$ and $u_{j'}$. Let $W_{\|}$ be the configuration with $x \| x'$ and $W_{\in}$ with $x \in x'$.

Assuming there are m elements along the chain U between $u_{j'}$ and u_i, then there are m more subsets in $W_{\|}$ than there are in $W_{\in}$. Following the argument as before, we must add in further elements to $W_{\in}$ in order to be able to compare the two configurations with one another. There are a number of different ways to add in new elements, always preserving the total number of subsets, and these different ways give different pairs of indexing elements for x and x', all of which are equally probable. But since we have $x \in x'$, all of these sets must preserve the condition that the indexing elements for x are contained in (or equal to) the indexing elements for x'. On the other hand, for $W_{\|}$ we now have fewer elements than $W_{\in}$ in the chain U, but more freedom to choose the indexing elements. All possibilities are open *except* the case where *both* indexing elements for x are contained in *both* indexing elements for x' (for that would imply that we must have $x \in x'$), or conversely, if both indexing elements for x' are contained in both indexing elements for x. Taken together, the number of possible configurations with $x \in x'$ is not greater than the number with $x \| x'$, and so the argument fails.

We see then that the presence of the second indexing chain, adding in so many further possible configurations in the case $x \in x'$, is essential for our argument. Having three or more indexing chains adds nothing, since they will only confirm the correlation between the ordering of the elements and that of their indexing elements. After all, the dimension is given by the least possible number of total orderings in a realizer.

A.5 Conclusion

From the few considerations dealt with here, it is obvious that our very natural way of defining probabilities in finite partially ordered sets will lead to structures which depart strongly from what might at first be expected. Rather than having a chaos of unordered sets, we see that chain-like structures which might interact with one another in orderly ways are probable.