CAUSAL SET THEORY
FROM
A STATISTICAL PERSPECTIVE

A REPORT SUBMITTED FOR THE CONTINUATION OF
REQUIREMENTS FOR A PROJECT DURING

BACHELOR OF SCIENCE
(RESEARCH)

BY

NIKSHAY CHUGH
UNDERGRADUATE PROGRAMME
INDIAN INSTITUTE OF SCIENCE

UNDER THE SUPERVISION OF

PROF. DEEPAK DHAR
INTERNATIONAL CENTER FOR THEORETICAL PHYSICS, BENGALURU

Acknowledgements

Everyone who listened to me blabber that spacetime may be discrete.

Abstract

In this report, I describe in detail my understanding of a possible discrete structure of spacetime.

We start with a partially ordered set of events on a flat Minkowski spacetime and construct spacetime coordinates for them. These are verified by explicit construction on a computer. Some analytic computations of chain lengths are then considered.

Next, we develop the theory of causality and causal spaces. We then discuss how causal sets may be an approach towards the discretization of spacetime. We talk about random processes in causal set theory.

Following this, we deal with the theory of the enumeration of partially ordered sets. Entropy is defined in such a system and computed for some constraints on the density of comparable pairs. We talk about first-order phase transitions in this system.

The next part of the project dealt with percolation, which I cover in my report. We simulate lattice animals by converting an existing FORTRAN code to Python. We use the ideas developed to conjure up percolating systems that may mimic spacetime. Some ideas are presented, and most are discarded for various reasons. We obtain nontrivial geometric problems that result in some physics.

Contents

Part I

Causal Sets

Chapter 1

Introduction

1.1 Events on a $(1+1)$-D Minkowski spacetime

Let's say we have a $(1+1)$-dimensional flat spacetime, with the Y-axis labelled by t and the X-axis labelled by x. We can represent the events in this spacetime as a set of points in the $t-x$ plane. The events are represented by the coordinates (t, x), where t is the time coordinate and x is the space coordinate.

Consider in this plane a diamond-shaped region with vertices at $(t, x) = (0, 0)$, $(T/2, T/2)$, $(-T/2, T/2)$ and $(0, T)$. In this region, we sample N points uniformly at random. From these points, we can construct a partially ordered set (poset) of events. The poset is constructed by defining a relation between the events based on their causal structure. Let there exist a set $\mathcal{C}$ such that

$$\mathcal{C} = \{\text{events in the diamond region}\} \tag{1.1}$$

Clearly, $|\mathcal{C}| = N$.

Equip this set $\mathcal{C}$ with a relation $\prec$ such that it is reflexive, antisymmetric, transitive, and locally finite.

1. **Reflexive:** For all $v_x \in \mathcal{C}$, $v_x \prec v_x$, $\forall x \in \{1, 2, ...N\}$.
2. **Antisymmetric:** For all $v_x, v_y \in \mathcal{C}$, if $v_x \prec v_y$ and $v_y \prec v_x$, then $v_x = v_y$.
3. **Transitive:** For all $v_x, v_y, v_z \in \mathcal{C}$, if $v_x \prec v_y$ and $v_y \prec v_z$, then $v_x \prec v_z$.
4. **Locally finite:** For all $v_x, v_y \in \mathcal{C}$, if $v_x \prec v_y$, then there exists a finite number of events v_z such that $v_x \prec v_z \prec v_y$.

Such a relation is called a **partial order** relation. The set $\mathcal{C}$ with the relation $\prec$ is called a **partially ordered set** (poset). In the context of spacetime, the poset represents the causal structure of the events in the diamond region. This is why we can call it a *causal set*.

We ask the following question:

Given the causal set $(\mathcal{C}, \prec)$, can we assign unique coordinates to the events in $\mathcal{C}$, if we know that we are embedding it in a flat $\mathbb{M}^2$ Minkowski spacetime?

An important point to note is that the events in $\mathcal{C}$ are sampled from a diamond-shaped region in $\mathbb{M}^2$. We, of course, know the answer to this question. The answer is **no**. The reason is that we can make infinitesimally small changes to the coordinates of the events in $\mathcal{C}$, and the causal structure will remain unchanged.

This makes us modify the question to the following:

Given the causal set $(\mathcal{C}, \prec)$, can we assign coordinates to the events in $\mathcal{C}$ upto some finite precision, if we know that we are embedding it in a flat $\mathbb{M}^2$ Minkowski spacetime?

This question is still not well-defined, because we can always make Lorentz transformations on the coordinates of the events in $\mathcal{C}$, and the causal structure will remain unchanged. We can change the origin to anywhere in $\mathbb{M}^2$ and the coordinates will change, but the causal structure will remain unchanged.

We further modify the question:

Given:

- *The causal set $(\mathcal{C}, \prec)$*
- *A minimal and a maximal element in $\mathcal{C}$, v_{min} and v_{max} respectively, which are fixed to be the events at the bottom and top of the diamond region respectively.*
- *The fact that we are embedding it in a flat $\mathbb{M}^2$ Minkowski spacetime.*

Can we assign coordinates to the events in $\mathcal{C}$ upto some finite precision?

Turns out that this is a well-defined question. The answer is **yes**. To investigate this, we need a notion of *finite precision*. Before we do that, let us first understand the causal structure of the events in $\mathcal{C}$.

1.1.1 Causal Sets

Right now, I will not motivate why we are interested in causal sets. It should suffice to say that they are a candidate for a theory of quantum gravity. The idea is to take the continuum limit of the causal set and obtain a smooth spacetime manifold. This is similar to how we take the continuum limit of a lattice model in statistical mechanics to obtain a smooth arena. There are of course many subtleties in this process. This will be the subject of subsequent chapters.

We have so far an object $(\mathcal{C}, \prec)$, which is a causal set. We can think of defining objects like

$$J^+(v_x) := \{v_y \in \mathcal{C} \mid v_x \prec v_y\} \tag{1.2}$$

$$J^-(v_x) := \{v_y \in \mathcal{C} \mid v_y \prec v_x\} \tag{1.3}$$

$$L_c(v_x) := J^+(v_x) \cup J^-(v_x) \tag{1.4}$$

$$[v_x, v_y] := \{v_z \in \mathcal{C} \mid v_x \prec v_z \prec v_y\} = J^+(v_x) \cap J^-(v_y) \tag{1.5}$$

where $J^+(v_x)$ is the future light cone of the event v_x, $J^-(v_x)$ is the past light cone of the event v_x, and $L_c(v_x)$ is the lightcone of the event v_x. The lightcone is the set of all events that are causally connected to the event v_x. $[v_x, v_y]$ is the set of all events that lie in the intersection of the future light cone of v_x and the past light cone of v_y. We call this the **Alexandrov interval** between the events v_x and v_y. This is a very important object in the theory of causal sets.

This $(\mathcal{C}, \prec)$ provides an intuitive picture for timelike connected events. *What about spacelike connected events?* We will return to this question momentarily.

Lorentz Invariance

We are solving our problem in

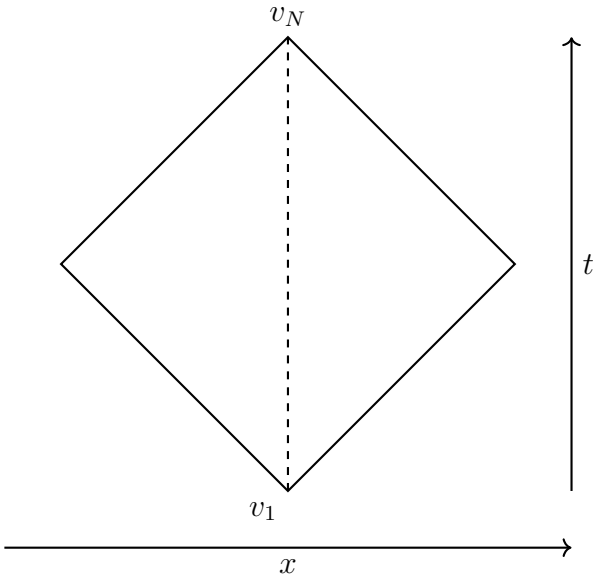


Figure 1.1: A diamond-shaped region in a $(1+1)$-D Minkowski spacetime

and we should think of this region as $\mathcal{C} = [v_1, v_N]$. To maintain Lorentz invariance, we need to figure out what this diamond looks like a boosted reference frame. This will result in *time dilation* and *length contraction* of the diamond region. We will see something like this: We must note that proper volume of the diamond region is invariant under Lorentz transformations. Thus, in this diagram, a uniform sampling of events in the diamond region is Lorentz invariant. A uniform sampling, i.e., a Poisson sampling, maintains Lorentz invariance. We will prove a

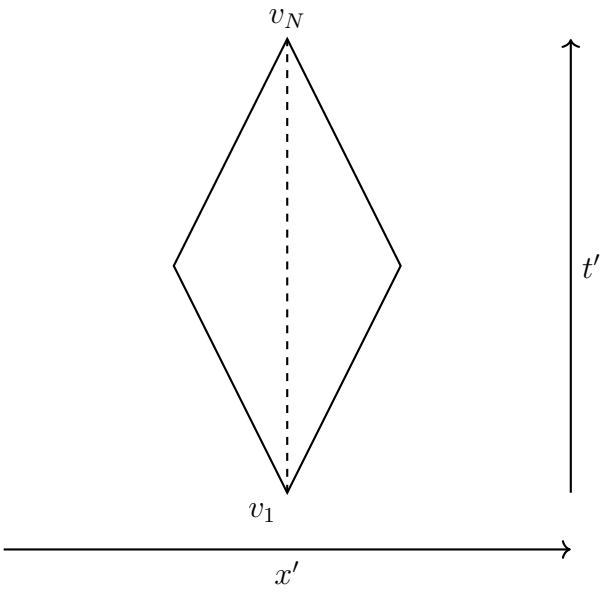


Figure 1.2: A boosted diamond-shaped region in a boosted $(1+1)$-D Minkowski spacetime

more general version of this result in later chapters.

Accesible points

For a point v_x in the diamond region, less number of events are accessible to it than the number of events accessible to the minimal element v_1. Hence, we should think of a *measure of proper*

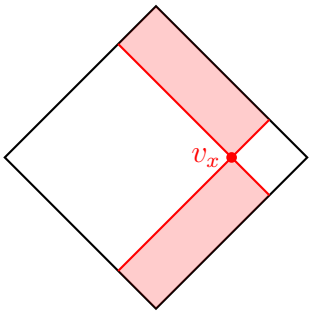


Figure 1.3: A point v_x and its light cone

time of each event to be related to the number of events accessible to it. All constructions of proper time in my report are based on this idea. Take the simplest case of choosing each step to be of unit proper time. By step, I mean the nearest future event of v_x in the causal set. This is a well-defined object.

An object that can be used to represent a causal set, or any general poset, is a **Hasse diagram**. A Hasse diagram is a graphical representation of a finite partially ordered set. The elements of the poset are represented by points, and the order relation is represented by lines connecting the points. The lines are drawn in such a way that if $a \prec b$, then a is drawn below

b. It is the minimal, transitively reduced representation of the poset. This is a directed acyclic graph (DAG). This ca further be used to define more concretely what we mean by proper time and steps.

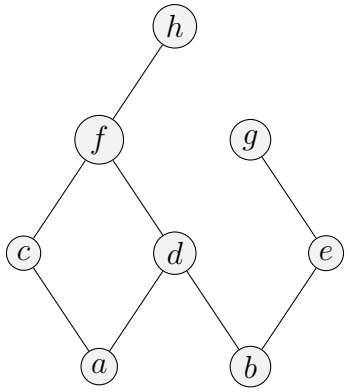


Figure 1.4: A random partially ordered set with 8 elements

The figure above shows a Hasse diagram of a random poset with 8 elements. The elements are represented by circles, and the order relation is represented by lines connecting the circles. Note that pairs (a, b), (c, d), (d, e), (d, g), etc. are not comparable. This means that there is no relation between them. In a Hasse diagram, we only draw links. Our current definition of proper time is $t_{\text{link}} = 1$. We will show why this is a good definition of proper time in due course.

1.1.2 Baby Steps

Let's gather our arsenal. We have a causal set $(\mathcal{C}, \prec)$ (which is a poset), a general idea of Lorentz invariance, a poor definition of proper time, and a way to represent the causal set as a Hasse diagram. Is this enough to solve our problem? No harm in trying.

Choose N events uniformy and randomly from the causal diamond. This is a Poisson sampling of events. Since a point is of measure zero with respect to a finite area, we can safely assume that the origin and the top of the diamond are not chosen. We could be more concrete and say that we choose N events uniformly from the diamond region and not from the edges, but it is important to note that these details wipe out in the large N limit. Now, choose the origin to the event v_1 and the top of the diamond to the event v_N. This is a well-defined choice. The set of events $\mathcal{C} = \{v_1, v_2, ...v_N\}$ is a finite set of events. Construct the relation $\prec$ on the events in $\mathcal{C}$ such that it follows from the causal structure of the events, and we have our causal set $(\mathcal{C}, \prec)$, and it is equal to the Alexandrov interval $[v_1, v_N]$. We will now be brave and forget where the locations of the events are. We want to reconstruct those using *some* method.

As a first step, let us identify the symmetries of our system, and how to get rid of them.

1. **Translation invariance:** The events in the causal set are invariant under translations. This means that we can shift the coordinates of the events by a constant amount, and

the causal structure will remain unchanged. This can be fixed by fixing the coordinates of the events v_1 as $(0,0)$. One can think of this as anchoring the events to a point in the spacetime.

2. **Rotational Invariance in Spatial Dimensions:** If this were a $(1+2)$-D or a $(1+3)$-D spacetime, we would have rotational invariance in the spatial dimensions. This here manifests as *Parity* invariance. We can fix this by choosing the direction of propagation of light to be along the positive x-axis.

3. **Lorentz Invariance:** As noted before the events in the causal set are invariant under homogenous Lorentz transformations. This means that we can boost the coordinates of the events by a constant amount, and the causal structure will remain unchanged. This can be fixed by choosing the point v_N to be at $(0,T)$, where T is the time coordinate of the event v_N. This is nothing but fixing the time coordinate of the event v_N to be T. Any other choice for v_N will be equivalent to this upto a homogenous Lorentz transformation.

We can summarise this using a map $p : \mathcal{C} \to \mathbb{M}^2$ such that

$$p(v_1) = (0,0) \quad \text{and} \quad p(v_N) = (0,T) \tag{1.6}$$

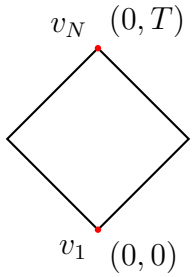


Figure 1.5: The diamond region with events v_1 and v_N

This map is a bijection between the events in the causal set and the points in the diamond region. The map p is not unique, but it is well-defined. We want to see to what extent does p vary.

1.1.3 Never Too Late To Look At The Time

We will adopt the easier terminology $p(v_x) = x^\mu$ for the coordinates of the event v_x. To embed points in $\mathbb{M}^d$ ($d = 2$ in our case),

$$v_x \prec v_y \implies x^\mu \prec y^\mu \equiv \left(x^0 \leq y^0 \quad \text{and} \quad \left(y^0 - x^0\right)^2 - \left(y^1 - x^1\right)^2 \geq 0\right) \tag{1.7}$$

for the metric $\eta_{\mu\nu} = \text{diag}(1, -1)$. Timelike vectors have a positive norm and their proper time is given by

$$\tau(y^\mu, x^\mu) = \sqrt{\langle y^\mu - x^\mu \mid y^\mu - x^\mu \rangle} \tag{1.8}$$

where $\langle \cdot | \cdot \rangle$ is the contraction in the Minkowski spacetime. This should be interpreted as

$$\tau(y^\mu, x^\mu) = \sqrt{(y^\mu - x^\mu)(y_\mu - x_\mu)} \tag{1.9}$$

For our system, there are a total of N points in the diamond region. The volume of this region is $\frac{T^2}{2}$. Hence, the density of events in the diamond region is given by

$$\rho = \frac{N}{\frac{T^2}{2}} = \frac{2N}{T^2} \tag{1.10}$$

For a general spatial dimension d, we can extend this to

$$c_d \cdot T^d \cdot \rho = N \tag{1.11}$$

where c_d is a constant that depends on the spatial dimension $d-1$. For our present purposes, the value of c_d is irrelevant for dimensions other than 2. It can be easily computed as the twice the volume of a right cone of height $T/2$, which turns out to be $c_2 = 1/2$. This method of thinking of volumes of cones is very useful in the theory of lightcones, as we will see in the final chapters.

We define another measure of proper time, based on the Alexandrov interval

$$\tau(v_x, v_y) = \left(\frac{|[v_x, v_y]|}{c_d \cdot \rho} \right)^{1/d} \tag{1.12}$$

We have two definitions of proper time now, both of these applicable to events that are timelike connected. Note that we have not yet defined a proper distance yet. The first definition of proper time is defined for links only, so we extend it to non-links as well. To do that, we need to recall a thought-experiment of Special Relativity.

Geodesics

Connecting two timelike connected events x and y ($x \prec y$) in a Minkowski spacetime is a straight line with speed less than or equal to the speed of light. This is a geodesic. This geodesic *maximises* the proper time between the two events. We can prove this almost trivially.

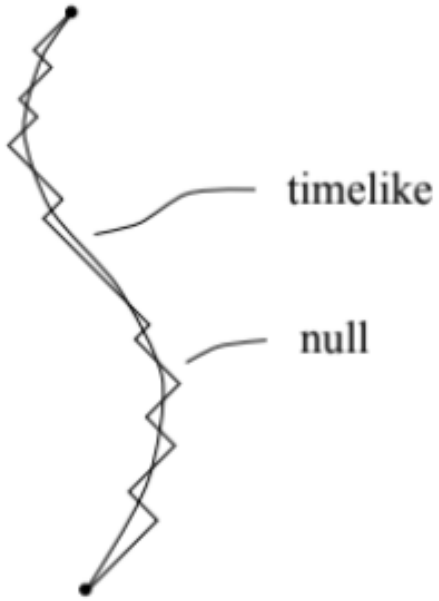


Figure 1.6: A geodesic in a $(1+1)$-D Minkowski spacetime

An appropriate relativistic action for the geodesic extremises the proper time between two events. We can say with surity that there exists a path between two events that minmises the proper time. This is the path taken by light rays. The above diagram neatly sums up the situation.

Two varities of proper time (for $v_x \prec v_y$)

1. **Inspired by Geodesics:** The proper time between two events v_x and v_y is given by the length of maximal chain between them. This is the length of the longest chain between the two events. This is a well-defined object, and it is equal to the number of links in the chain.

$$C(v_x, v_y) = \{\text{chains between } v_x \text{ and } v_y\} \tag{1.13}$$

$$\tau_G(v_x, v_y) = \max_{\zeta \in C(v_x, v_y)} \{|\zeta|\} \tag{1.14}$$

2. **Inspired by Invariant Intervals:** The proper time between two events v_x and v_y is given by the *length* of the Alexandrov interval between them, defined by equation Eq. (1.12).

$$\tau_A(v_x, v_y) = \left(\frac{|[v_x, v_y]|}{c_d \cdot \rho} \right)^{1/d} \tag{1.15}$$

So far, there is no reason to believe that these two definitions of proper time are equivalent, or even related. They are related in the continuum limit by geometric relations apparent from the metric. We have, so far, only heurestic reasoning to believe that these two definitions are even

related to proper time. One should note that $\tau_G(v_1, v_N) = \tau_A(v_1, v_N) = T$.

1.1.4 A Little Bit Of Geometry - Part 1

For three elements $x, y, z \in \mathbb{M}^d$, we have, from the following triangle

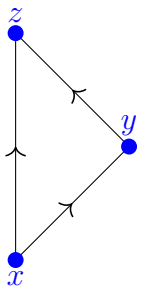


Figure 1.7: A triangle in a $(1+1)$-D Minkowski spacetime

that results in the following equations

$$\langle z-x|z-x\rangle = \langle (z-y)+(y-x)|(z-y)+(y-x)\rangle \tag{1.16}$$
$$= \langle z-y|z-y\rangle + \langle y-x|y-x\rangle + 2\langle z-y|y-x\rangle \tag{1.17}$$

which implies that

$$\langle z-y|y-x\rangle = \frac{1}{2}\Big(\langle z-x|z-x\rangle - \langle z-y|z-y\rangle - \langle y-x|y-x\rangle\Big) \tag{1.18}$$

for any three events $x, y, z \in \mathbb{M}^d$.

We develop an analog of this on the causal set.

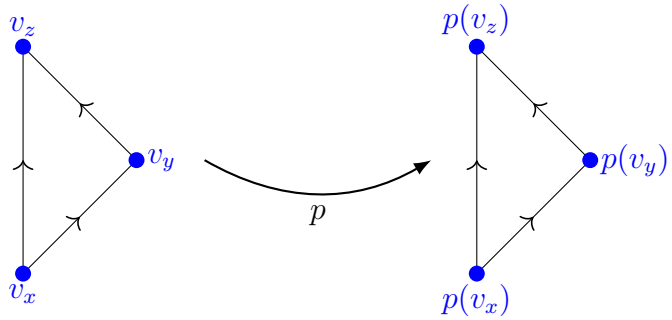


Figure 1.8: A triangle in $(1+1)$-D Minkowski spacetime and its image under a map p

It makes sense to think of the following figures

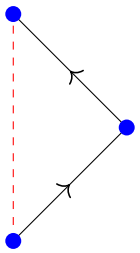


Figure 1.9: Triangle in a Causal Set

and the associated causal contraction

$$\langle v_z - v_x | v_z - v_x \rangle_{\mathcal{C}} := \frac{1}{2}\Big[\tau(v_x, v_z)^2 - \tau(v_x, v_y)^2 - \tau(v_y, v_z)^2\Big] \tag{1.19}$$

and two other triangles obtained by choosing the other two edges of the triangle. This holds only for the case when $v_x \prec v_y \prec v_z$.

We are now in a position to define the time coordinates of the events in the causal set. Since all events in the causal set are timelike connected to the minimal element v_1, we can define the time coordinates of the events in the causal set as

$$t(v_x) := \langle v_x - v_1 | v_N - v_1 \rangle_{\mathcal{C}} \times \frac{1}{T} \tag{1.20}$$

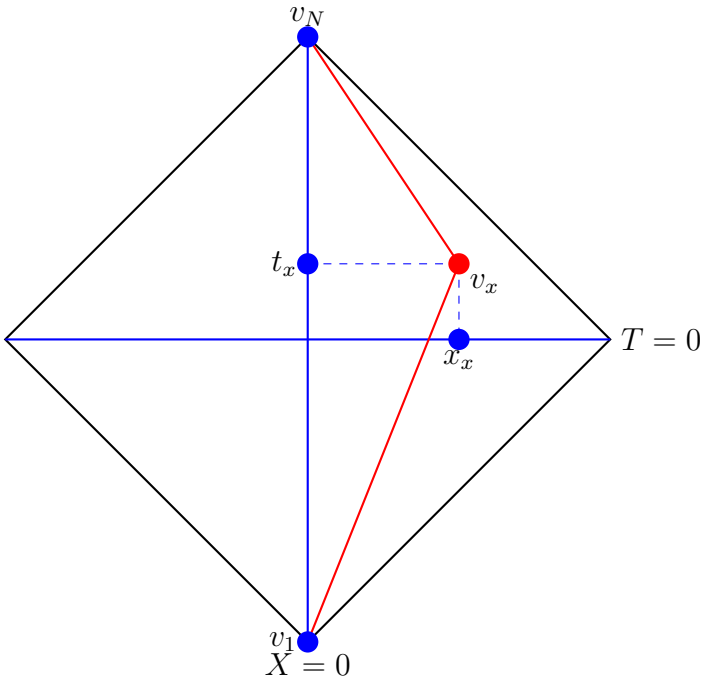
where T is the time coordinate of the event v_N which comes into play to implement dimensional consistency.

$$t(v_x) = \frac{T}{2} + \frac{\tau(v_1, v_x)^2 - \tau(v_x, v_N)^2}{2T} \tag{1.21}$$

for timelike connected events. There can be arbitray number of steps between v_1 and v_x and between v_x and v_N. Due to the existence of two varities of proper time intervals, we can see that there are two types of time coordinates t_G and t_A.

1.1.5 Pythagoras Is Not Dead (Yet)

Consider the following picture:



Using this picture, we can assign x coordinates to the events in the causal set. To do this, solve either of

$$\left(\frac{T}{2} - t(v_x)\right)^2 + x(v_x)^2 = -\tau(v_x, v_N)^2 \tag{1.22}$$

$$\left(\frac{T}{2} + t(v_x)\right)^2 + x(v_x)^2 = -\tau(v_1, v_x)^2 \tag{1.23}$$

We see here that there is not one, not two, but four ways to assign the same event an x coordinate. This is because the event v_x can be connected to the events v_1 and v_N. This is obviously just an apparent issude since these equations are equivalent (see yourself, its simple). So there remain two ways to assign the event v_x an x coordinate, exactly the number of parity states. On average, in the large N limit, the distribution of events will be symmetric in density about the $X = 0$ line. Hence, for the two point that we find that have the same x coordinate magnitude, we can place one on the left and one on the right. This seems like a reasonable choice, and is subject to verification by numerical simulations. The algorithm to do this has been presented in the code below.

1.1.6 Antichains

Another method to construct x coordinates is by writing down antichains. An antichain is a set of elements in a poset such that no two elements are comparable. In other words, for any two elements a and b in the antichain, neither $a \prec b$ nor $b \prec a$. A formal description requires the

construction of

$$\mathfrak{p}(v_x, v_y) := \left\{ v_z \in J^-(v_x) \cap J^+(v_y) \mid \max\left(t(v_z)\right) \right\} \tag{1.24}$$

$$\mathfrak{f}(v_x, v_y) := \left\{ v_z \in J^+(v_x) \cap J^-(v_y) \mid \min\left(t(v_z)\right) \right\} \tag{1.25}$$

where $\mathfrak{p}$ is the common past with maximum time coordinate and $\mathfrak{f}$ is the common future with minimum time coordinate. From two events v_x and v_y, we have found elements $\mathfrak{p}(v_x, v_y)$ and $\mathfrak{f}(v_x, v_y)$ that are *definitely* timelike connected, with $\mathfrak{p}(v_x, v_y) \prec \mathfrak{f}(v_x, v_y)$, even if v_x and v_y are not. This is a very useful construction.

Using this, we can construct the following object

$$l(v_x, v_y) := \tau(\mathfrak{p}(v_x, v_y), \mathfrak{f}(v_x, v_y)) \tag{1.26}$$

which gives the proper time interval when the events v_x and v_y are timelike connected, and gives a measure of spacelike separation when they are not. This shows that we can find maximally spacelike separated events.

If we consider the set of unrelated events to be $\mathcal{C}_s$, we can construct an order on it by using τ_G to be our proper time. We can then employ all our developed technology to determine the space coordinates of the events in $\mathcal{C}_s$. This completes our construction of the coordinates of the events in the causal set.

1.1.7 Further Ideas

Maximal Chain Assignment of Coordinates

Uniformly distribute the maximal chain on the line from v_1 to v_N. We expect there to be a unique maximal chain. If there isn't then we choose one of them. Due to symmetry, we can expect that the distribution of events will be symmetric about the $X = 0$ line. In the next step, we assign time coordinate to the next pair of maximal chains. We keep doing this till we reach the two maximally spacelike separated events.

Since the maximal chain length may not be unique, as discussed above, we could think of allowing some tolerance in the chain length.

In a flat sapcetime with d spatial dimensions, we would expect the number of maximal chains to be d. We are naturally led to think of the number of maximal chains as the number of spatial dimensions, and the next natural question is to ask if we given a causal set, can we determine what was the dimension of spacetime that it was embedded in. This is a very interesting question, and we will try to cover it in the next chapter.

Tolerance of Chain length

We can assign a tolerance of ϵ to the chain length. This is a measure of how much we are willing to deviate from the maximal chain at each step. This tolerance may be related to some aspect of the physics of the system. For example, in a quantum system, the tolerance may be related to the uncertainty principle.

Shape of Path

Let's say in the causal diamond, we have two paths P_1 and P_2 connecting the minimal and maximal element that are of the same length (number of links) but have different shapes. P_1 is nearly straight with small deviations from the straight line $X = 0$, while P_2 is a zig-zag path with large deviations from the straight line. In our most latest method to assign coordinates, we have not taken into account the shape of the path. *Which of these should we choose as the maximal chain?* We must note that there will be overlapping points in P_1 and P_2 due to the prescription of our problem. This problem can arise at each step of the construction.

Local substructure

In this construction, we cannot control what will happen if we encounter paths that look the following.

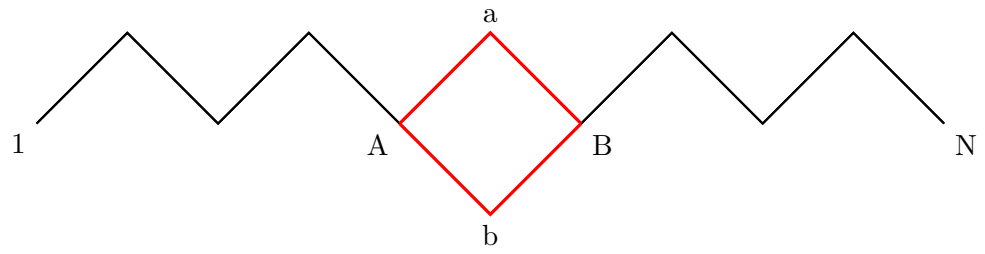


Figure 1.10: Zigzag path from 1 to N with a central plaquette (A, a, b, B).

In such a case, even though the paths are different, there is an inevitable intermixing of the events a and b when we assign coordinates. This feature is built into the theory.

Expected Region for each event

Given a certain element of $\mathcal{C}$, what is the expected region of the diamond that it will occupy over the course of multiples simulations? Fortunately, this is something that we can answer with a reasonable amount of confidence.

We can use the concept of **Voronoi tessellation** to answer this question. Given points $P_1, P_2, \ldots, P_n$ in a space, the Voronoi tessellation is a partitioning of the space into regions such

that each region contains all points that are closer to a given point than to any other point. In our case, we can think of the points P_i as the events in the causal set.

By Euler's formula, in a 2D space with no holes, $F + V - E = 2$, where F is the number of faces, V is the number of vertices, and E is the number of edges. In our case, each vertex opens into three edges (is seen from the Delauny triangulation) and each edge shares two vertices. Hence, we can write the equation as

$$2 + E = 2E/3 + F \tag{1.27}$$

and in the large E limit, we can ignore the 2 and write the equation as $E = 3F$. This means that the number of edges is three times the number of faces. The average number of edges per Voronoi cell is 6, since each edge shares two faces. So, we expect the regions occupied by each event to be hexagonal in the large density limit. *What about the size of the hexagons?* This question is simple to anwer since we can divide the area of the diamond by the number of events in the causal set. The area of the diamond is $T^2/2$ and the number of events is N. Hence, the area of each hexagon is given by the density of events in the diamond region.

1.1.8 Numerical Results

Here, I sketch out the methods in detail that I used to generate the numerical results. I have constrcuted a class called 'CausalDiamond' that implements the methods to generate the causal set and assign coordinates to the events in the causal set. The code is written in Python. It is available on GitHub.

We have first sprinkled points in a causal diamond, then we have computed the proper time (both varities) and used this to assign coordinates using both the methods. There are interesting physical observations to be made using both methods.

We used the following examples to generate the results. The outputs are printed below the inputs.

```
# Sprinkle 1000 points in the causal diamond
N = 100
causal_diamond = CausalDiamond(N)
# Plot the points in the causal set
points = causal_diamond.sprinkle_points(N)
causal_set = causal_diamond.build_causal_set()

causal_diamond.plot_causal_set(points)
```

Listing 1.1: Sprinkling points and plotting

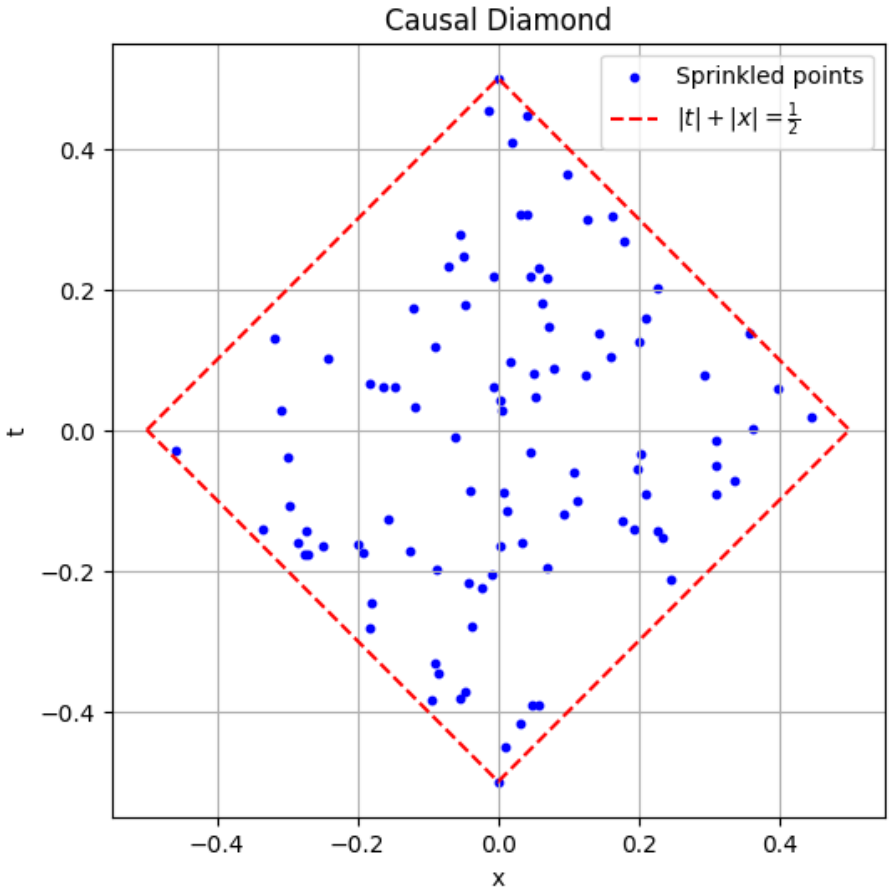


Figure 1.11: Causal diamond with 100 points.

Then we generate plots of the assigned coordinates obtained by the geodesic and Alexandrov interval proper time prescriptions.

```
# Example usage
N = 100
cd = CausalDiamond(N)
points = cd.sprinkle_points(N)
causal_set = cd.build_causal_set()

cd.plot('G')
cd.plot('A')
```

Listing 1.2: Plotting the assigned coordinates

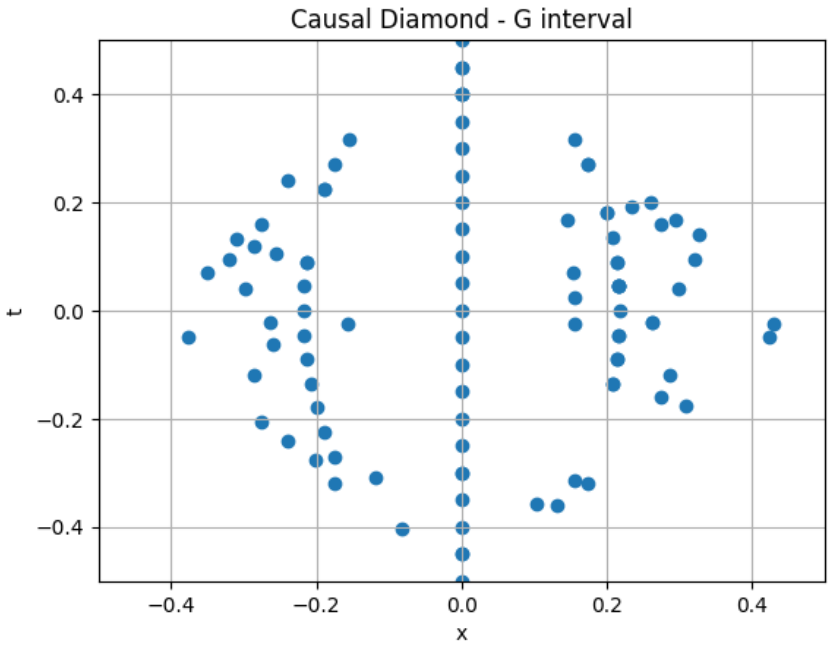


Figure 1.12: Assigned coordinates using t_G.

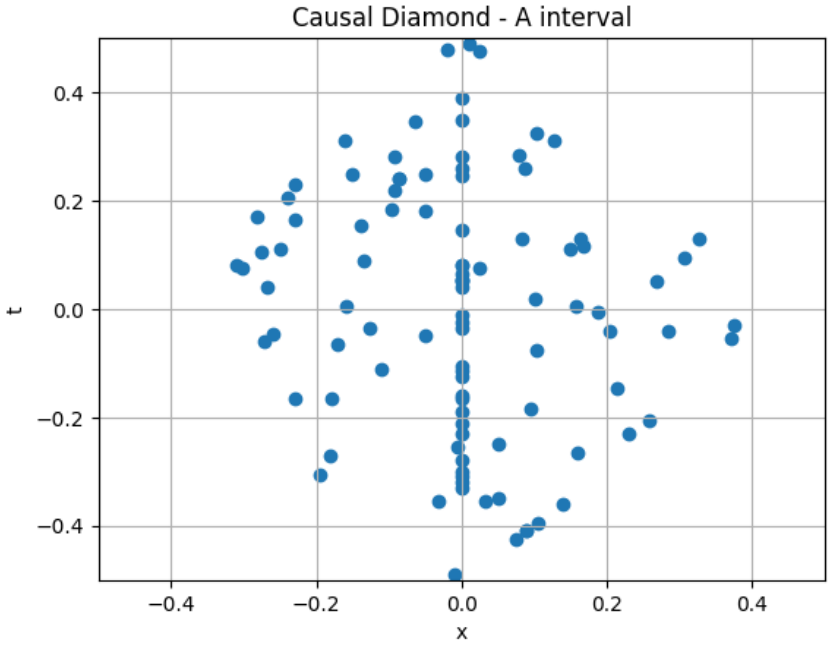


Figure 1.13: Assigned coordinates using t_A.

There is *much* to say here. Generally, advocates of CST as a theory of quantum gravity use the t_G prescription to talk about actions (to deal with dynamics). The t_A prescription is used to talk about the geometry of the causal set. Here, we see that the t_G prescription results in the kind of assignment that we talked about when we were discussing 1.1.7. We did not expect that these would line up so spectacularly. The t_A prescription results in a more uniform distribution

of events.

It is imperative to note that the assignment of the maximal chain in the t_G prescription is (roughly) unique in all realizations. Although the number of steps in it may vary, the assigned coordinates are all equally spaced on the $X = 0$ line. Such is not the case for the t_A prescription. In fact, in most cases, there is a high chance that there are not many equally spaced events on the $X = 0$ line for the t_A prescription. However, in both cases, the length of the maximal chain is almost the same $\sim (2N)^{1/2}$.

We can now talk about *dimension estimators*. There can be multiple ways to estimate the dimension of a causal set, given it has been sampled from a flat spacetime.

- **Length of Maximal Chain:** For a d dimensional spacetime, the length of the maximal chain is expected to be $\sim (N/c_d\rho)^{1/d}$, where c_d is a constant that depends on the dimension of the spacetime and ρ is the density of events in the causal set. We will later analytically show that we approach this limit for large N from below.
- **Distribution of Chains of some length vs Length:** These are characteristic to the dimension you are living in. One can check this on their own since it is out of the main focus of this project. This estimator is called the *Myrheim-Meyer dimension*.
- **Alexandrov proper time between minimal and maximal elements:** This is different from the length of the maximal chain, just like t_G and t_A are different. It is often also called the *Midpoint-scaling dimension*.

The final thing we do with this problem, at least in a numerical sense, is to compute the distance distribution of the events in the causal set. This is done by computing the Euclidean distance between the assigned coordinates and the original coordinates, and then averaging over many realizations.

```
# Comparative distance distributions
N = 100
cd = CausalDiamond(N)
cd.plot_comparative_distance_distributions(n=200, bins=250, bin_range=(0, 1))
```

Listing 1.3: Computing the distance distribution

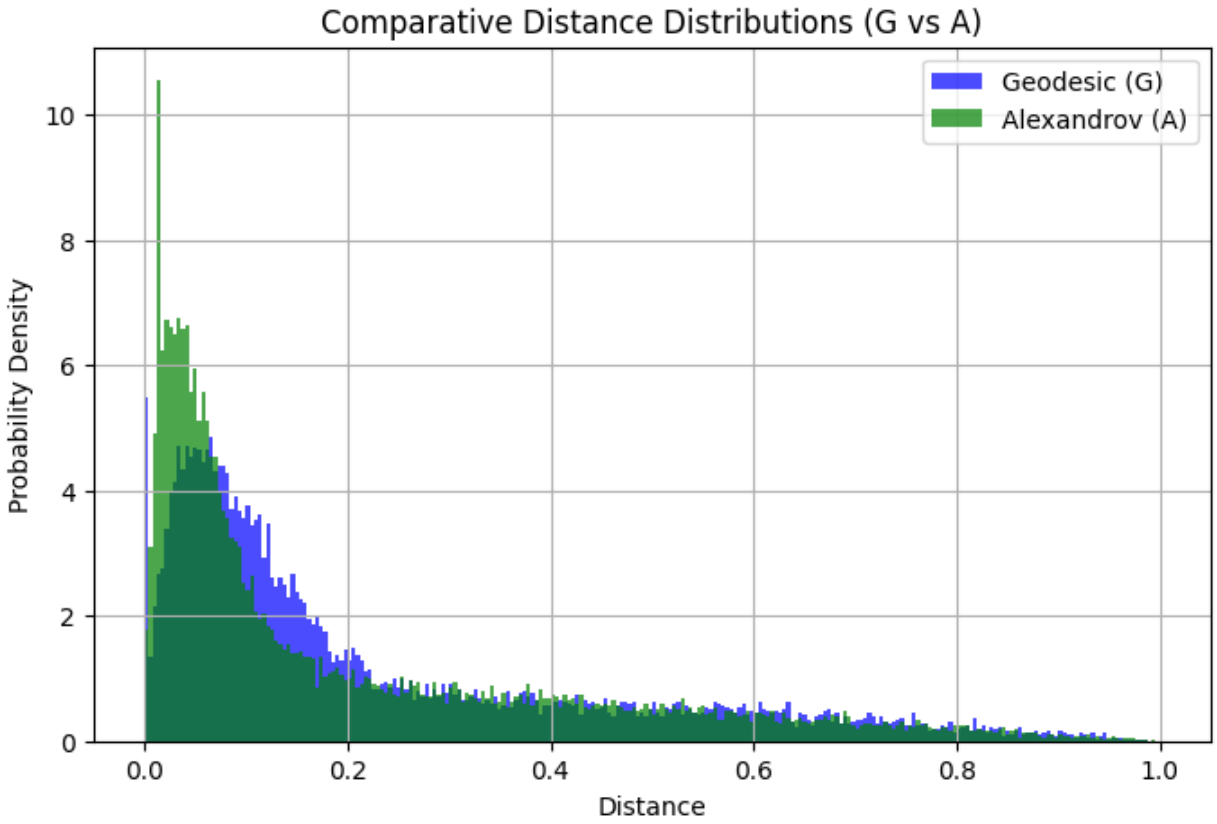


Figure 1.14: Distance distribution of the assigned coordinates.

Though we do not yet know the exact nature of the probability distributions, we can observe that the mean for the t_A prescription is less than the mean for the t_G prescription, and the variance too. From here, we can take the general lesson that the t_A prescription is *better* than the t_G prescription.

We should also note that the t_G prescription is not that bad an approximation. We hold tightly to this idea because this idea is crucial in three aspects of the geometrodynamic theory: curvature is more easily described by the counting of links, a (reasonably good, though non-local) action based on links has been a nice step towards path integrals in literature, and sequential growth of the causal set is most easily described by the counting of links.

1.2 Analytic Results in $(1+1)$-D Spacetime

In this section, we will compute some important, instructive, but trivial results in $1+1$ dimensions. We set out with the aim to compute by hand the expectation value of the length of the maximal chain in a causal set sampled from a diamond in $\mathbb{M}^2$. To do this, we will draw the Hasse diagrams of causal sets with $N = 1, 2, 3$ and 4 events. We will see that this calculation is nontrivial for $N \geq 4$ and it would be easier to let a computer do it. We will show that the $(2N)^{1/2}$ scaling is approached strictly from below.

In general, when we construct posets, we encounter a layered structure that we will describe briefly in this section and in detail in the next chapter. Only a limited fraction of these are derivable my random sampling from a manifold. How do we choose them? We construct a plausible weight and show using explicit computation that it is a bad choice.

1.2.1 Computation of Probabilities of Small Hasse Diagrams

Consider a causal diamond with side length normalized to 1, so that it has an area of 1. For this section we define the length of the maximal chain to be the number of links in the longest chain plus one.

1. **N=1:** The only possible causal set is the trivial one with a single event. The length of the maximal chain is 1. The probability of this causal set is 1.

2. **N=2:** There are two possible causal sets: one with a single link and one with no links. The length of the maximal chain is 2 for the first case and 1 for the second case. The probability of the first case is 1/2 and the probability of the second case is 1/2. This results in an average length of the maximal chain of 3/2.

   This can be seen in the following way too:

Figure 1.15: Hasse diagrams for N=2

By symmetry, they must be equally likely. The average length of the maximal chain is given by 3/2. In order to make life complicated for now, and easier for later, we can think of the following method. Consider a point chosen at a random coordinate (x, y) in the causal diamond, with the orientation of the axes as shown in the figure 1.16.

$$\mathbb{P}(\text{other point lies in past or future}) = \int_0^1 \int_0^1 \Big[(xy) + (1-x)(1-y)\Big] \mathrm{d}x\mathrm{d}y \tag{1.28}$$

$$= 1/2 = \mathbb{P}(l = 2) \tag{1.29}$$

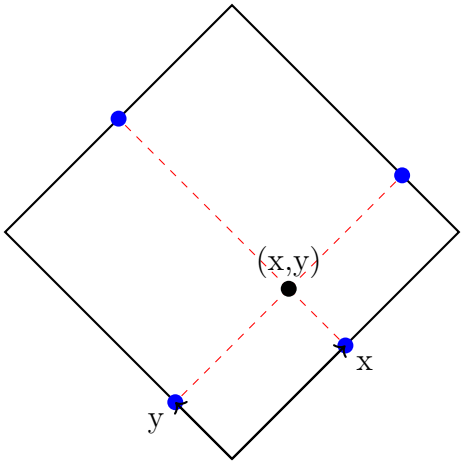


Figure 1.16: A point in the causal diamond

Similarly,

$$\mathbb{P}(\text{other point does not lie in past and future}) = \int_0^1 \int_0^1 \Big[(1-x)(y) + (x)(1-y)\Big] \mathrm{d}x\mathrm{d}y \tag{1.30}$$

$$= 1/2 = \mathbb{P}(l = 1) \tag{1.31}$$

Labelling the areas by I, II, III, IV in anticlockwise order, we can write this as

$$\mathbb{P}(l = 2) = \int_{\text{all}(x,y)} \mathbb{P}(B \in \text{I} \cup \text{III} \mid A = (x, y)) \mathrm{d}x\mathrm{d}y \tag{1.32}$$

$$\mathbb{P}(l = 1) = \int_{\text{all}(x,y)} \mathbb{P}(B \in \text{II} \cup \text{IV} \mid A = (x, y)) \mathrm{d}x\mathrm{d}y \tag{1.33}$$

3. **N=3:** The possible posets are

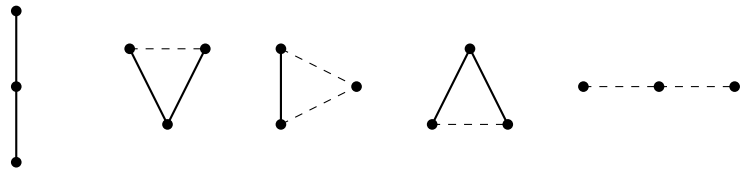


Figure 1.17: Hasse diagrams for N=3

By $X - T$ symmetry we can say that $\mathbb{P}(l = 3) = \mathbb{P}(l = 1)$ and $\mathbb{P}(l = 2u) = \mathbb{P}(l = 2d)$. In fact, if we look a little deeper, we will see that there is a symmetry betweem presence and absence of links. We see that $\mathbb{P}(l = 2u) \times 2 = \mathbb{P}(l = 2)$. This means that

$$2\mathbb{P}(l = 2) + 2\mathbb{P}(l = 3) = 1 \tag{1.34}$$

There may be more subtle ways to proceed here, but I will compute $\mathbb{P}(l = 3)$ explicitly instead of exploring them. From exactly the same method as equations 1.32, we can partition our diamond into 9 regions, and use Mathematica to solve the integral. This results in $\mathbb{P}(l = 3) = 1/6$ and $\mathbb{P}(l = 2) = 1/3$. Hence, the average length of the maximal chain is 2..

4. **N=4:** The possible posets are

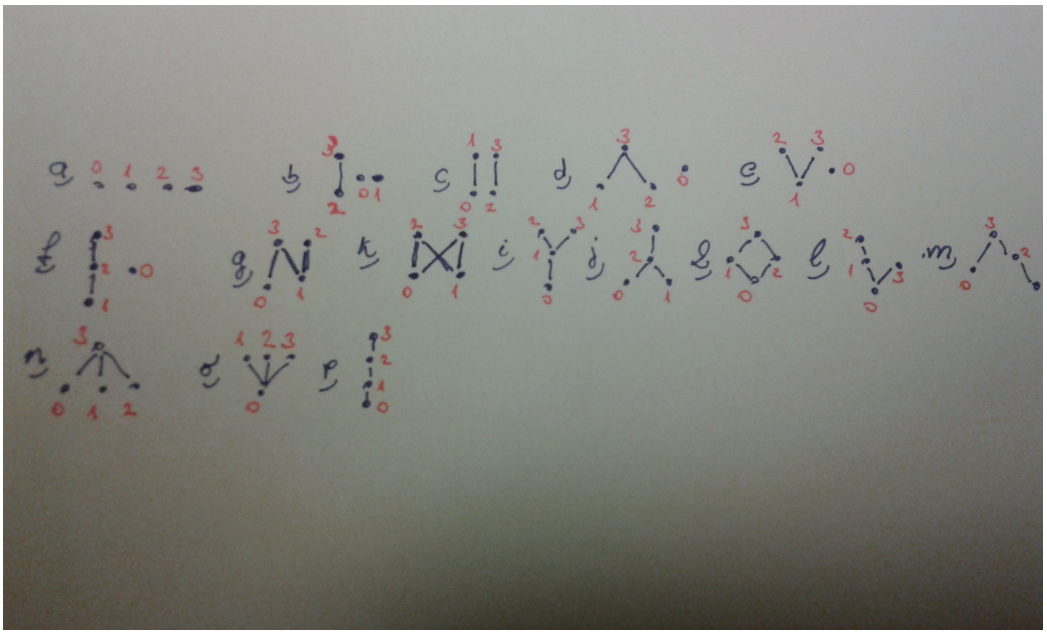


Figure 1.18: Hasse diagrams for N=4

It is clear that the number of possible posets grows rapidly with the number of elements. To be complete, we note that

$$\mathbb{P}(\text{configuration}) = \frac{\text{Number of permutations that result in this configuration}}{\text{Total number of permutations}} \tag{1.35}$$

Is there a way to enumerate partial orders on N element sets? Turns out, due to Kleitman and Rothschild, there is a way to do this approximately.

1.2.2 Layers Of Bricks Make A House

Let there exist a set P of N elements. This set is endowed with a partial order $\prec$. This is no arbitrary set. It has the following structure

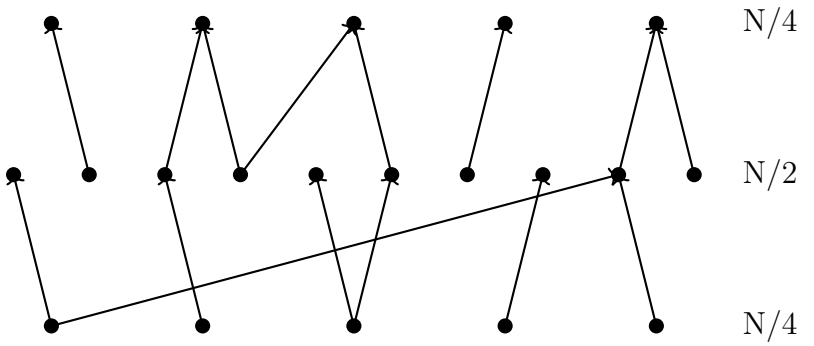


Figure 1.19: Three-layer structure of the poset

Basically, it can be divided into two layers of $N/4$ elements and one layer of $N/2$ elements. Only adjacent layers are connected and other relations are understood by transitivity. The number of ways to connect the elements in the top layer to the middle layer is $2^{N/4 \cdot N/2}$. The number of ways to connect the elements in the middle layer to the bottom layer is $2^{N/2 \cdot N/4}$. The total number of ways to connect these layers is $2^{N^2/4}$. This is the number of distinct possible partial orders on N element sets with the given structure.

Now, take δ of these nodes from the top to the bottom layer. We get the following structure.

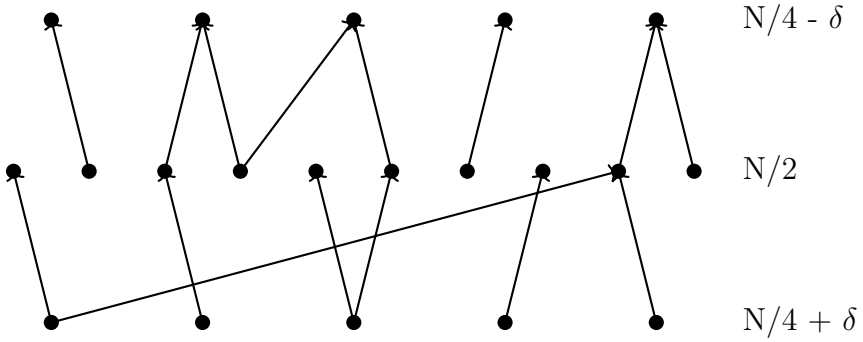


Figure 1.20: Modified three-layer structure of the poset: 1

In this, we have

$$2^{(N/4-\delta)(N/2)} \cdot 2^{(N/2)(N/4+\delta)} = 2^{N^2/4} \tag{1.36}$$

partial orders. This means that no matter what δ is, we have the same number of possible partial orders. Let's mix this up to the following. This gives us $2^{N^2/4-\delta^2}$ partial orders. However, $\max \delta = N/4$, at which we have the minimum number of partial orders as $2^{N^2/4-N^2/16} = 2^{3N^2/8}$. The factor of $1/4$ and $3/8$ will be crucial when we construct the entropy of these sets.

From Boltzmann's definition of entropy, we can say that $S \propto \log(\Omega)$, where Ω is the number of distinct partial orders. This means that in this three-layer structure, entropy behaves as $S \propto N^2$.

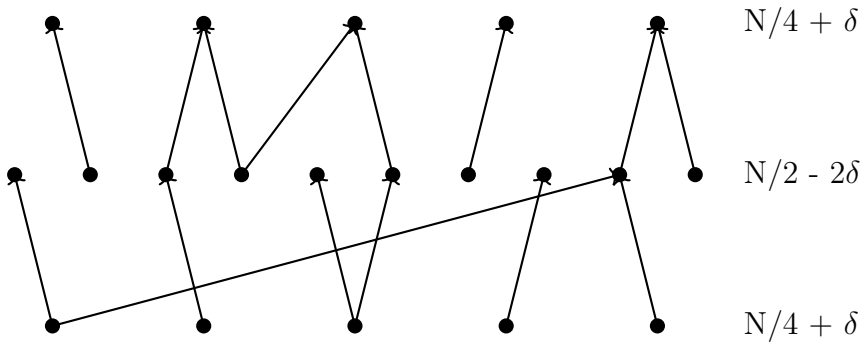


Figure 1.21: Modified three-layer structure of the poset: 2

This is not extensive. In fact it is way worse than we expect, because most of the posets that can be construct with finitely many layers are quadratic in N. We need infinitely many layers to get extensive entropy.

Since we can't demand that partial orders be the fundamental building blocks of spacetime, given the above argument, we desperately try to find a way to hold onto the fact that when we sample a causal set from a manifold, we get a partial order, and this has extensive entropy. This implies that *manifold-like* posets are of the order $\exp N$. This is expected from a Poisson sampling too.

If we suppose that the fundamental geometrodynamical theory *is* derivable from the distinct partial orders that can be described on a set of events, then to get manifolds as the expected result, we would need to assign some weight to each configuration of partial orders. This weight must be such that the expected poset that is obtained from sampling from the set of posets is manifold-like. To correct for the overcounting due to the abundance of finitely-many-layered posets, we need to introduce statistical weights to each configuration such that *uniformly embeddable* posets are more likely to be sampled.

This idea is a central question in Causal Set Theory. Once we have a working definition of how to obtain the weights for a poset with N elements, we could immediately generalise it to a grand canonical ensemble and obtain dynamics. This is a very hard problem, and no one has solved it exactly yet. In the next subsection, we try to construct a possible solution and in the subsection after that, we show how we are wrong.

1.2.3 A Possible Solution

For a configuration $\mathcal{C}$ of points and their partial order relations, we need to define

$$\mathbb{P}(\mathcal{C}) = \frac{\exp(-\beta E_{\mathcal{C}})}{\mathcal{Z}} \tag{1.37}$$

where $\mathcal{Z}$ is the partition function given by $\mathcal{Z} = \sum_{\mathcal{C}} \exp(-\beta E_{\mathcal{C}})$ and $E_{\mathcal{C}}$ is the energy of the configuration $\mathcal{C}$. We need to conjure up a sensible defintion of $E_{\mathcal{C}}$.

A flat (1+1)-D spacetime has a well-defined relation between $\tau_G(\min, \max)$ and $\tau_A(\min, \max)$ for the minimal and maximal element in a causal diamond. For the minimal point at $(0, 0)$ and the maximal point at $(0, T)$, we know that the proper time is T and the volume of the causal interval is $T^2/2$. We work with the natural objects in this picture; the causal volume between timelike connected points a and b (V_{ab}) and the proper time between them (τ_{ab}). In the discrete picture, these look like $|[a, b]|$ and $\tau_G(a, b)$ respectively. We can try a function of the form

$$E_{\mathcal{C}} = \sum_{(a,b) \in \mathcal{C}}^{\text{all timelike pairs}} f(|[a, b]|, \tau_G(a, b)) \tag{1.38}$$

where f is a function that we need to define. We want the energy to be nearly zero or negative if we are near a flat Lorentzian manifold, and large and positive otherwise.

The simplest approaximation is to take the energy as a quadratic monomial, just like in a harmonic oscillator. This will give us a Gaussian distribution of posets centered at those that are *flat*. Adopting the same symbols as continuum (i.e. $V_{ab} \equiv |[a, b]|$ and $\tau_{ab} \equiv \tau_G(a, b)$), we can write

$$f(V_{ab}, \tau_{ab}) := \frac{(V_{ab} - c_d \cdot \tau_{ab}^d)^2}{[\sigma(\tau_{ab})]^2} \tag{1.39}$$

working in $d - 1$ spatial dimensions, where c_d is the same constant as before, and we can choose a fluctuating geometry based on the function $\sigma(\tau_{ab})$. Recall $c_2 = 1/2$.

How to choose the fluctuations?

In a Poisson sprinkling,

$$\mathbb{P}(n \text{ points in a volume } V) = \frac{(\rho V)^n}{n!} e^{-\rho V} \tag{1.40}$$

where ρ is the density of points in the volume V. Points are randomly chosen at a fixed density and the total number of points sampled is propertional to the volume. In such a system

$$\langle n \rangle = \rho V \tag{1.41}$$

$$\langle n^2 \rangle_c = \rho V = \text{var}(n) \tag{1.42}$$

We can thus choose the variance to be ρV_{ab}. Our expression for f then becomes

$$f(V_{ab}, \tau_{ab}) := \frac{1}{\rho} \frac{(V_{ab} - c_d \cdot \tau_{ab}^d)^2}{V_{ab}} \tag{1.43}$$

Note that because density shows up like this, we can redefine β to incorporate this as $\beta = \kappa\rho$. Infinite temperature then corresponds to zero density and zero temperature corresponds to infinite density, all determined by a scale factor κ.

1.2.4 Numerical Results

Upon running this simulation for different values of κ, we find that the causal set is reduced from roughly $\sqrt{2N}$ layers to a finite number of layers. It results in three layered structures in the large steps limit.

[NOT SURE OF THE INTERPRETATION ASK PROF!!!]

Chapter 2

Gravity

Hurling apples at people sitting under trees is rarely a nice a idea, unless they are physicists. What happens when you do throw an apple at a physicist? The apple falls to the Earth. Then the physicist deduces that the Earth falls towards the apple too. A couple of centeries later, it is more generally accepted that neither is falling due to the other, but both are freely falling in some background curvature, and this curvature is determined by the freely falling masses. By this time, the apple has usually hit the physicist on the head, and they are not too happy about it.

2.1 The Math Behind The Physics

In the general theory of relativity, the physics is diffeomorphism invariant. This means that any physical observable does not depend on the coordinates we use to describe the manifold. This is a very strong condition. This invariance manifests locally on the manifold as invariance under the action of the homogenous Lorentz group. Due to the manifold having non-zero curvature, we cannot demand global Lorentz invariance. However, using the connection and parallel transport, we can extend these ideas.

There was a general concensus that Lorentz invariance and Causality are related, but until Zeeman, there was no concrete proof. In the following subsection, we sketch out his derivation, and sort out some other stuff. First, we show that the group of chronological automorphisms is isomorphic to the group of inhomogeneous Lorentz transformations and dilations on $\mathbb{M}^d$. Second, we show how to construct abstract causal spaces axiomatically from the causal poset and the chronological poset. Third, we state the HKMM theorem. Fourth, we show that causal bijections imply chronological bijections. Finally, we collect some results pertaining to the topology of the causal set. In the following discussion, I assume that details about differential geometry are known to the reader.

We need to build up some preliminary terminology for the discussion. Consider an event-set

denoted by M. This set has a lot of freedom, particularly it can be either discrete or it can be a manifold. Depending on the use below, we will give M some concrete status. On this set, we define a *chronological* relation $\ll$ and a *causal* relation $\prec$. Both of these prescribe a partial ordering. Further, we define $x \to y$ if $x \prec y$ but not $x \ll y$.

To make our discussion somewhat concrete, let us take the event-set to be a smooth manifold X equipped with a Lorentzian metric. We identify the tangent-space T_x of x in X and define the null-cone N_x as the subset of the tangent-space on which the vectors result in zero metric length. We can easily now define vectors in T_x to be timlike, spacelike, or null, according to them lying inside the null-cone, outside the null-cone, or on the null-cone respectively. The removal of the zero vector and the exterior of N_x decomposes T_x into two disjoint sets *future-directed* and *past-directed*. The properties of the metric allows us to do perform this locally. We call a smooth Lorentzian manifold X *future-distinguishing* and *past-distinguishing* if the future-directed and past-directed vectors are disjoint at each point in X. We now only deal with such structures. This is our *first restriction*.

From our *first restriction*, we can define *orientation* of non-spacelike paths in X, saying that a path is *positive* if the tangent-vectors everywhere are future-directed. The relation xTy (resp. xNy) means that there exists a positive timelike (resp. non-spacelike) path from x to y. These relations satisfy, for all $x, y, z \in X$:

1. if xTy, then xNy;
2. if xNy and yNz, then xNz;
3. if xTy and yNz, or if xNy and yTz, then xTz.

We can now define the relations as

1. $x \prec y$ iff xNy or $x = y$;
2. $x \ll y$ iff xTy;
3. $x \to y$ iff either xNy and not xTy, or $x = y$.

We demand our *second restriction* to be that the manifold is *causal* which translates that xNx for no point x in X, or in simpler terms that there exist no closed non-spacelike curves. This construction automatically excludes compact manifolds. An interesting caveat is covered in Kronheimer and Penrose, where they tackle the question of whether every causal space, taken here as manifold, is locally Minkowskian. The answer is NO.

In general, we call the tuple $(X, \prec, \ll, \to)$ a *causal space* if X is a set and $\prec, \ll, \to$ are relations on X such that for all x, y, z in X

1. $x \prec x$;

2. if $x \prec y$ and $y \prec z$, then $x \prec z$;

3. if $x \prec y$ and $y \prec x$, then $x = y$;

4. not $x \ll x$;

5. if $x \ll y$ then $x \prec y$;

6. if $x \prec y$ and $y \ll z$, then $x \ll z$;

7. if $x \ll y$ and $y \prec z$, then $x \ll z$;

8. $x \to y$ iff $x \prec y$ and not $x \ll y$;

The set X is called the *underlying set*, and the relations are called *causality* ($\prec$), *chronology* ($\ll$), and *horismos* ($\to$). An immediate consequence of the axioms is that, whereas the causality is a reflexive and the chronology an anti-reflexive partial ordering, the horismos is a horismotic relation.

2.1.1 The Early Bird *Always* Gets The Worm

Theorem 1. *The group of chronological automorphisms* ($\mathrm{Aut}_{\ll}(M)$) *on* $\mathbb{M}^d$ *for* $d > 2$ *is isomorphic to the group of inhomogeneous Lorentz transformations and dilations* ($\mathrm{Lor}(M)$) *on it.*

Proof. A chronological automorphism is a bijection $f : M \to M$ where M is the underlying set such that both f and f^{-1} preserve the chronology. We make no assumption of f being linear or continuous.

It is trivial to note that $\mathrm{Lor}(\mathbb{M}^d) \subseteqq \mathrm{Aut}_{\ll}(\mathbb{M}^d)$ since the generators of the former preserve chronology. The converse is true only for dimensions greater than 2. If we were working with $\mathbb{M}^2$, then the following counterexample shows why this is not true.

The characteristic two-form is

$$Q(x) = x_0^2 - x_1^2, \quad x = (x_0, x_1) \in \mathbb{M}^2 \tag{2.1}$$

Choose a new coordinates system

$$y_0 = x_0 - x_1, \quad y_1 = x_0 + x_1 \tag{2.2}$$

Let f_0 and f_1 be two arbitrary nonlinear orientation-preserving homeomorphisms of the real line onto itself. Define $f : \mathbb{M}^2 \to \mathbb{M}^2$ as

$$f(x_0, x_1) = (f_0(y_0), f_1(y_1)) \tag{2.3}$$

Then f is a chronological automorphism of $\mathbb{M}^2$ that is not a Lorentz transformation.

We need five lemmas to prove our theorem. I will state these but not prove them. Refer to Zeeman's paper for the proofs.

Lemma 1. *Let $f : M \to M$ be a function that is a one-to-one mapping. Then, f and f^{-1} preserve chronology iff f and f^{-1} preserve horismos.*

Lemma 2. *A chronological automorphism maps light rays to light rays.*

Lemma 3. *A chronological automorphism maps parallel light rays to parallel light rays.*

So far, everything works for two dimensions as well. We will need some subtlety in the proof of the following lemma which excludes the two-dimensional case.

Lemma 4. *A chronological automorphism maps each light ray linearly.*

Lemma 5. *A chronological automorphism maps parallel equal intervals on light rays to parallel equal intervals.*

Given a chronological automorphism f, we can show that it is linear. To do this, first compose it with the necessary translation such that the origin is invariant. Then, decompose a point x in the manifold as a linear combination of d light rays from the origin. From Lemma 4, we know that f maps light rays linearly. This means that the image of x under f is a linear combination of the images of the light rays. This means that f is linear.

Now due to Lemma 1, we know that f preserves the null-cone. Since the metric on the null-cone is zero, this would still be left invariant if f were multiplied by some scalar.

This proves that apart from translations and dilations, f is a linear Lorentz transformation. Therefore, $f \in \mathrm{Lor}(\mathbb{M}^d)$. □

2.1.2 Constructing Abstract Causal Spaces

In Kronheimer and Penrose, the authors construct three types of causal spaces $\mathfrak{A}$, $\mathfrak{B}$, and $\mathfrak{C}$ from the relations $\to$, $\ll$, and $\prec$ respectively. We use table 2.1 to work on the following. Due credit is given to Kronheimer and Penrose, since most of this subsection is almost directly taken from their work. Here, we outline the construction of causal spaces from the relations $\to$, $\ll$, and $\prec$, and the related results on the topology of such spaces.

Definition 1. *A causal space is called* regular *if for x_1, x_2, y_1, and y_2 distinct points in the space and $x_i \to y_j$ for each i and j, then the horismos orders x_1 and x_2 iff it orders y_1 and y_2.*

This is not the only definition of regularity and we construct another equivalent one below. The idea behind this definition is to map null geodesics of a flat Minkowski manifold to straight lines. To think more concretely, consider a point x_1 in $\mathbb{M}^2$ and a pair of points y_1 and y_2 lying on the future null geodesics of x_1. If both of these lie on the same null geodesic, then they are related by a horismos. If they lie on different null geodesics, then they are not related by a

horismos. The regularity condition demands that if x_1 is related to y_1 and y_2 by a horismos, and if y_1 and y_2 are related by a horismos, then in order for x_2 to be related to y_1 and y_2 by a horismos, it must be related to x_1 by a horismos. This is a very strong condition, and it is not satisfied by all causal spaces. Simply speaking, all of x_1, x_2, y_1, and y_2 must lie on the same null geodesic if $x_i \to y_j$ for each i and j along with y_1 and y_2 being related by a horismos.

Lemma 6. *Let x, y, z be points in a causal space. If $x \prec y \prec z$ and $x \to z$, then $x \to y \to z$.*

The above lemma allows us to reinterpret the definition of regularity. A causal space is regular if for x_1, x_2, y_1, and y_2 distinct points in the space and $x_i \to y_j$ for each i and j, then $x_1||x_2$ iff $y_1||y_2$.

Symbol	Name	Definition / Interpretation
$J^+(a)$	Causal future of a	The set $\{x : a \prec x\}$; all points causally in the future of a
$I^+(a)$	Chronological future of a	The set $\{x : a \ll x\}$; all points strictly chronologically in the future of a
$C^+(a)$	Null future of a	The set $\{x : a \to x\} = J^+(a) \setminus I^+(a)$; points causally but not chronologically in the future
$J^-(a)$	Causal past of a	The set $\{x : x \prec a\}$; all points causally in the past of a
$I^-(a)$	Chronological past of a	The set $\{x : x \ll a\}$; all points strictly chronologically in the past of a
$C^-(a)$	Null past of a	The set $\{x : x \to a\} = J^-(a) \setminus I^-(a)$; points causally but not chronologically in the past
$[p, q]$	Causal interval	The set $\{x : p \prec x \prec q\}$; all points causally between p and q
$\langle p, q\rangle$	Chronological interval	The set $\{x : p \ll x \ll q\}$; all points strictly chronologically between p and q
$x\|\|y$	Spacelike	x and y are spacelike related; $\neg(x \prec y \vee y \prec x)$

Table 2.1: Futures and pasts of a point a in a causal space.

There is an obvious duality in this system, everything holds if we replace $\prec$, $\ll$, and $\to$ with their inverses, and replace "future" with "past".

Definition 2. *Let $\to$ and $\prec$ be two relations on a set X. We say that $\to$ is a horismos* compatible *(resp.* regularly compatible*) with the causality $\prec$ if for some relation $\ll$ the quadruple $(X, \prec, \ll, \to)$ is a (resp. regular) causal space. The set of all horismoi (regularly) compatible with the causality $\prec$ for the specified X is will be written as*

$$\{(\text{reg.})\ \text{hor} \mid \text{cau}\ \prec\}$$

We can define other such sets by mixing hor, cau, chr *in the two positions similarly.*

We now define an important concept that forms the basis of the Hawking-King-McCarthy-Malament theorem, an important result that motivates the construction of causal sets.

Definition 3. *Let X be a set and $\ll$ an anti-reflexive partial ordering on X. We shall call* $\ll$ future-reflecting *if $I^-(x) \subset I^-(y)$ whenever $I^+(x) \supset I^+(y)$. (Dually, $\ll$ is* past-reflecting *if $I^+(x) \supset I^+(y)$ whenever $I^-(x) \subset I^-(y)$. These are independent properties.)*

We call $\ll$ weakly distinguishing *if $x = y$ whenever both $I^+(x) = I^+(y)$ and $I^-(x) = I^-(y)$,* future-distinguishing *if $x = y$ whenever $I^+(x) = I^+(y)$.*

We call $\ll$ full *if it satisfies both the following condition and its dual: For each $x \in X$ there exists p such that $p \ll x$; and, if $p_1 \ll x$ and $p_2 \ll x$, there exists q such that $p_1 \ll q$, $p_2 \ll q$, and $q \ll x$.*

These defintions of future-reflecting, weakly distinguishin and future-distinguishing are important because they are useful in defining rigorously the *first restriction* but they do not automatically imply the *second restriction*. We need to impose the *second restriction* separately.

A short note on the topology of causal spaces is in order, though I have not done any work on it yet. Based on the definitions of the causal and chronological relations (not the horismos), we can construct two distinct topologies on the causal space. Since the causal space is not necessarily a manifold, these are not necessarily the topologies of a manifold. Moreover, this topological causal space may not be Hausdorff. The following table summarizes the two topologies that can be constructed on a causal space.

Symbol	Name	Definition / Interpretation
$\mathcal{T}^*$	Alexandrov Topology	Constructed from the chronological relation $\ll$. Defines open sets as those containing all points that are chronologically after or before any point in the set, i.e. $\forall x \in X$, $I^-(x) \in \mathcal{T}^*$ and $I^+(x) \in \mathcal{T}^*$.
$\mathcal{T}^+$	Causal Topology	Constructed from the causal relation $\prec$. Defines closed sets as those containing all points that are causally before any point in the set, i.e. $\forall x \in X$, $J^-(x)$ is closed in the topology $\mathcal{T}^+$.
$\mathcal{T}^{\text{man}}$	Manifold Topology	Defined only if the causal space is a manifold. Defines open sets as those that are in the standard topology of the manifold.

Table 2.2: Topologies associated with a causal space.

Many lemmas and theorems can be proved about these topologies. We stress the most relevant ones here.

Lemma 7. *The following statements for a causal space X which is also a manifold are equivalent:*

1. $\mathcal{T}^* = \mathcal{T}^{\text{man}}$;

2. *the topological space $(X, \mathcal{T}^*)$ is Hausdorff;*

3. *if each point of $I^-(a)$ precedes each point of $I^+(b)$, then b cannot strictly precede a;*

4. *if $u \to v$ and each point of $I^-(v)$ chronologically precedes each point of $I^+(u)$, then $u = v$;*

5. *each point of X possesses an arbitrarily small open neighborhood, with respect to $\mathcal{T}^{\text{man}}$, whose natural causal structure as a submanifold of X coincides with its structure as a causal subspace of X.*

and another one,

Lemma 8. *If $(X, \mathcal{T}^*)$ is Hausdorff and $\ll$ is full, then $\ll$ is future- and past-distinguishing.*

We can now proceed to construct abstract causal spaces from the relations $\to$, $\ll$, and $\prec$.

$\mathfrak{A}$ From the Horismos:
Let X be a set and $\to$ a horismos on X. We define the relations $\prec^{\mathfrak{A}}$ and $\ll^{\mathfrak{A}}$ as follows:

(a) $x \prec^{\mathfrak{A}} y$ iff there exists a finite sequence $(u_1, u_2, \ldots, u_n)$ such that $x = u_1 \to u_2 \to \ldots \to u_n = y$;

(b) $x \ll^{\mathfrak{A}} y$ iff $x \prec^{\mathfrak{A}} y$ and not $x \to y$.

We have obtained $(X, \prec^{\mathfrak{A}}, \ll^{\mathfrak{A}}, \to)$ as a causal space $\mathfrak{A}$. So, $\prec^{\mathfrak{A}} \in \{\text{cau} \mid \text{hor} \to\}$ and $\ll^{\mathfrak{A}} \in \{\text{chr} \mid \text{hor} \to\}$. Since, for any $\prec$ in $\{\text{cau} \mid \text{hor} \to\}$, $x \prec y$ whenever $x \prec^{\mathfrak{A}} y$, it follows that $\prec^{\mathfrak{A}}$ is the intersection of all possible causal relations defined over the horismos $\to$. This means that $\prec^{\mathfrak{A}}$ is the *weakest* causal relation compatible with the horismos $\to$. Similarly, $\ll^{\mathfrak{A}}$ is the *weakest* chronological relation compatible with the horismos $\to$.

$\mathfrak{B}$ From the Chronology:
Let X be a set and $\ll$ an anti-reflexive partial ordering on X. We can define two types of causal spaces from this relation; one which is weakly distinguishing ($\mathfrak{B}_W$) and one which is future-reflecting ($\mathfrak{B}_+$). An anti-reflexive partial ordering is future-reflecting and weakly distinguishing iff it is future-reflecting and future-distinguishing.

We define the relations $\prec^{\mathfrak{B}_W}$ and $\to^{\mathfrak{B}_W}$ as follows:

(a) $x \prec^{\mathfrak{B}_W} y$ iff $I^+(x) \supset I^+(y)$ and $I^-(x) \subset I^-(y)$;

(b) $x \to^{\mathfrak{B}_W} y$ iff $x \prec^{\mathfrak{B}_W} y$ and not $x \ll y$.

If $\prec \in \{\text{cau} \mid \text{chr} \ll\}$, then $x \prec^{\mathfrak{B}_W} y$ whenever $x \prec y$. On the other hand if $x \prec^{\mathfrak{B}_W} y$, we can find a $\prec$ in $\{\text{cau} \mid \text{chr} \ll\}$ such that $x \prec y$. This establishes that $\prec^{\mathfrak{B}_W}$ is the strongest causal relation compatible with the chronology $\ll$. Similarly, $\to^{\mathfrak{B}_W}$ is the strongest horismotic relation compatible with the chronology $\ll$. This argument holds only if $\ll$ is weakly distinguishing.

We can also define the relations $\prec^{\mathfrak{B}+}$ and $\rightarrow^{\mathfrak{B}+}$ similar to above, with the modification that $x \prec^{\mathfrak{B}+} y$ iff $I^+(x) \supset I^+(y)$. This will result in a causal space only if the chronology is future-reflecting and future-distinguishing.

$\mathfrak{C}$ **From the Causality:**
Let X be a set and $\prec$ a reflexive partial ordering on X. We can define a causal space $\mathfrak{C}$ from this relation by defining the relations $\ll^{\mathfrak{C}}$ and $\rightarrow^{\mathfrak{C}}$ as follows:

(a) $x \rightarrow^{\mathfrak{C}} y$ iff

i. $x \prec y$, and

ii. $\prec$ linearly orders $[u, v]$ whenever $[u, v]$ is a proper subset of $[x, y]$;

(b) $x \ll^{\mathfrak{C}} y$ iff $x \prec y$ and not $x \rightarrow^{\mathfrak{C}} y$.

Then $(X, \prec, \ll^{\mathfrak{C}}, \rightarrow^{\mathfrak{C}})$ is a causal space $\mathfrak{C}$. We should note that $\prec$ is itself a horismos with ϕ a chronology compatible with this horismos. This implies that neither both chronology and horismos are weakest compatible with the causality, nor both are strongest compatible with the causality. It can be shown that

$$\rightarrow^{\mathfrak{C}} = \cup \{\text{reg. hor} \mid \text{cau} \prec\} \quad \text{and} \quad \ll^{\mathfrak{C}} = \cap \{\text{reg. chr} \mid \text{cau} \prec\}$$

I leave the proof of this to Appendix A.

Various causal spaces can such be constructed, and some examples are presented in Appendix A.

2.1.3 The Hawking-King-McCarthy-Malament Theorem

Definition 4. *A point p in a spacetime is said to be* strongly causal *if every neighborhood of p contains a subneighborhood such that no causal curve intersects it more than once. All the events in a* strongly causal spacetime *are strongly causal.*

Theorem 2. *If a map $f_b : M_1 \rightarrow M_2$ is a chronology preserving bijection between two d-dimensional strongly causal spacetimes which are both future and past distinguishing, then these spacetimes are conformally isometric when $d > 2$.*

Proofs of these theorems are presented in Appendix B.

Levichev showed that a causal bijection implies a chronological bijection. This means if we have the causal structure of a d-dimensional spacetime, then we can find its metric upto a conformal factor. Parrikar and Surya showed that causal spaces and their topology contain information about the dimension of the spacetime.

2.1.4 Motivation for Causal Sets

We treat spacetime as a manifold. The standard topology of a manifold is locally Euclidean, but we know as an experimentally verified fact that spacetime is Lorentzian. We generally do not indulge in the topology of a Lorentzian manifold as an object in its own right; we use the local standard topology of an Euclidean manifold and assign a Lorentzian metric to it. This works well but it does not capture the actual local topology of the Lorentzian manifold.

The previous discussion has shown that the causal structure of a spacetime is very rich. Inspired by the words of Finkelstein, we can say that the causal structure of a spacetime is 9/10 of the spacetime. The remaining 1/10 is the volume element, which fixes the conformal factor of the metric. The central proposal of *Causal Set Theory* is to consider a causal space which is a set, and try to fix the volume element by the cardinality of the set. Having defined just a causal relation on this set can prescribe a $\mathfrak{C}$-space. This would result in many benefits. First, a continuum spacetime would be a *full* extension of this set. Second, figuring out the topology of a discrete set is a much more tractable problem than that of a continuum. Third, since the set is discrete, we can use statistical methods to study how the metric emerges from the causal structure. Fourth, a discrete set of events may help us define path integrals in a sensible way.

In the following discussions, we worry about three properties of a causal set that are important for the emergence of a Lorentzian manifold from it. The first question that we will tackle is *how to find a causal set that will look like a manifold if it were full?* The second question is *how to maintain Lorentz invariance in a causal set?* And the third question is *are there other types of causal sets that are not Lorentzian manifolds?* We will tackle the first two questions in this chapter and beat the third question to a pulp in the next chapter.

2.2 Random Processes

According to section 1.1, given a causal set C with N (finite number) elements, we want to embed it in a Lorentzian manifold M such that the causal structure of C is preserved in M. We saw that the causal structure may be preserved upto small deviations in the spacetime coordinates. Let's twist this a bit. If instead we took the collection of all possible sets of points in M that admit the same causal structure as C, then we can think of the points in C as being randomly chosen from this collection. We can then think of this as a random process.

Upon hearing the word "random", one immediately thinks of probability distributions. *What is the probability distribution of this sampling?* This is a very good question, because the obvious, and coincidentally the correct, answer is to sample uniformly from the set of all points on the manifold. However, we can construct posets of N elements that look nothing like a Lorentzian manifold. Consider for example the trivial chain of N elements, where each element is causally related to the next one. This is a causal set, but it doesn't look like a finite subset of a Lorentzian manifold. One could choose the time-axis to be the chain and since this is infinite, we cannot

fix the volume element. This is a problem, because we want to fix the volume element by the cardinality of the causal set.

Recall the map $p : \mathcal{C} \to M$ from chapter 1. We call this an *embedding* of the causal set $\mathcal{C}$ in the manifold M. Clearly, we want the embedding to be a bijection. We also want it to be *uniform* over the spacetime volume measure. A causal set is said to approximate a spacetime $\mathcal{C} \sim (M, g)$ at *density* $\rho_c = V_c^{-1}$ if there exists a *faithful embedding* $p : \mathcal{C} \to M$ such that $p(\mathcal{C})$ is uniformly distributed over the spacetime volume measure of M at density ρ_c. We call the volume V_c the *discretization scale*.

2.2.1 How to Sample a Causal Set?

It seems painfully obvious that a regular lattice would not work as a causal set. The reason is that any regular lattice breaks continuous symmetries of a manifold to discrete symmetries. This is not what we want. We want to preserve the continuous symmetries of the manifold. Although any sprinkling of points would break the continuous symmetries due to the number of points being finite, we can hope to preserve it over an ensemble of sprinklings. In Christ, Friedberg, and Lee's works on random lattice field theory, it is shown that a uniform sprinkling via a Poisson process preserves on average the continuous symmetries of a Euclidean manifold. The expected behavior under any symmetry transformation is identical to the original process, even though individual realizations break the symmetry.

From section 1.2.3, we note that the standard deviation over the average number of points in a Poisson process is $1/\sqrt{\rho V}$ for sprinkling density ρ and volume V. This goes to zero for a large volume. This means that for a large volume, we can expect absolutely no deviation from the average number of points in a Poisson process. This is a very good property, because it means that we can expect the sprinkling to be uniform over the spacetime volume measure. We can ask what is the probability of finding macroscopic voids in such a sprinkling. In an infinitely large volume of a Poisson process, the probability of finding a void of any size is one.

This would imply that CST predicts that the universe is either infinite, abundantly filled with voids, and that we live in a sector which surprisingly doesn't, or that the universe is finite, in which case the predicted probability for a nuclear voide of size $10^{-60}m^4$ is of the order $10^{84} \times 10^{168} \times e^{-10^{72}}$ which is a very small number (Dowker et al 2004), or that we live in an infinite universe in which on average we do not find any voids. For our present purposes, all of these are equally likely. [ASK PROF ABOUT THIS!!!]

2.2.2 Emergence of Non-Locality

There is a major caveat in uniformly sampling a causal set via a Poisson process. Consider an element a in a causal set $\mathcal{C}$. The probability that any other event x in $\mathcal{C}$ to be a link is equal to the probability that the Alexandrov interval between a and x contains only a and x.

We can compute this exactly for a d-dimensional Minkowski spacetime. Consider the point a to be at the origin. Let's say the coordinates of x are $(t, x_1, x_2, \ldots, x_{d-1}) = (x_0, x_1, x_2, \ldots, x_{d-1})$. The the Alexandrov interval is given by intersection of light cones.

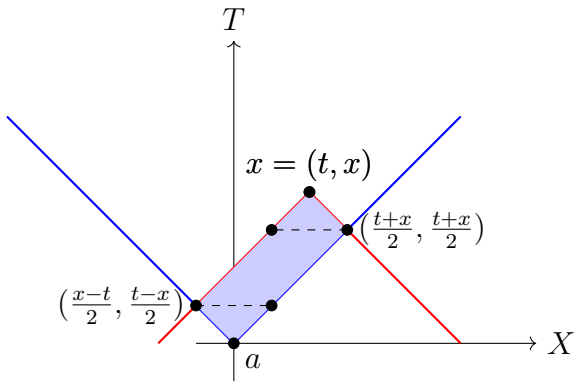


Figure 2.1: Alexandrov interval in 2D Minkowski spacetime. The point x is at (t, x) and the point a is at the origin. The Alexandrov interval is the blue shaded region. The red lines are the past light cone of x. The dashed lines are the lines of constant x.

In this, we can compute the volume of the Alexandrov interval

$$V(t,x) = 2 \times \left| \frac{1}{2} \frac{(t-x)}{2} (x-t) \right| + |x \times (t-x)| \tag{2.4}$$

$$= \frac{t^2 - x^2}{2} \tag{2.5}$$

If instead of $\pi/4$ the opening angle of the light cone was θ, and (t, x) is in the lightcone, then the volume would be

$$V(t,x) = \frac{t^2 \tan(\theta)}{2} - \frac{x^2}{2\tan(\theta)} \tag{2.6}$$

This will be useful in Part II. However, this sort of computation is not possible in higher dimensions, because the Alexandrov interval is not a simple geometric object which breaks into two cones and a parallelogramoid. In higher dimensions, we can construct a *modified Lorentz transformation* to solve this.

Let's say we are in d spatial dimensions. The modified light cones opens angle θ. So, we need to fix

$$|\mathbf{r}| = \pm \tan(\theta) t \tag{2.7}$$

on the modified light cone. We have taken $c = 1$ for simplicity. We now write a linear transfor-

mation $\Lambda_{\text{mod}} : (t, |\mathtt{r}|) \to (t', |\mathtt{r}'|)$ as

$$|\mathtt{r}'| = A|\mathtt{r}| + Bt \tag{2.8}$$
$$t' = C|\mathtt{r}| + Dt \tag{2.9}$$

and we need to have

$$|\mathtt{r}'| = \pm \tan(\theta) t' \tag{2.10}$$

whenever $(t, |\mathtt{r}|)$ is on the modified light cone, with the appropriate sign. This gives us

$$\frac{A \cdot \tan(\theta) + B}{C \cdot \tan(\theta) + D} = \tan(\theta) \tag{2.11}$$
$$\frac{-A \cdot \tan(\theta) + B}{-C \cdot \tan(\theta) + D} = -\tan(\theta) \tag{2.12}$$

This implies that $A = D$ and $B = C \cdot \tan^2(\theta)$. Let's say $B = -A\omega$, so $C = -A\omega \cot^2(\theta)$.

We want the determinant of the transformation to be 1, so we have

$$\det(\Lambda_{\text{mod}}) = 1 \tag{2.13}$$
$$A^2 - A^2\omega^2 \cot^2(\theta) = 1 \tag{2.14}$$
$$\frac{1}{\sqrt{1 - \omega^2 \cot^2(\theta)}} = A \tag{2.15}$$

We have the transformation

$$|\mathtt{r}'| = \frac{|\mathtt{r}| - \omega t}{\sqrt{1 - \omega^2 \cot^2(\theta)}} \tag{2.16}$$
$$t' = \frac{t - \omega \cot^2(\theta)|\mathtt{r}|}{\sqrt{1 - \omega^2 \cot^2(\theta)}} \tag{2.17}$$

We now want to change our coordinates such that $|\mathtt{r}'| = 0$. This implies $|\mathtt{r}| = \omega t$. We can find the volume as two double cones of height $t'/2$ and base angle θ.

$$\text{Vol}_d(t, |\mathtt{r}|) = \frac{\pi^{d/2}}{(d+1) \cdot \Gamma(\frac{d}{2} + 1)} \cdot \frac{t'^{d+1}}{2^d} \cdot \tan^d(\theta) \tag{2.18}$$

where

$$t' = t\sqrt{1 - \omega^2 \cot^2(\theta)} \tag{2.19}$$
$$= \left(t^2 - |\mathtt{r}|^2 \cot^2(\theta)\right)^{1/2} \tag{2.20}$$

which implies that the volume of the Alexandrov interval in d spatial dimensions is

$$\mathrm{Vol}_d(t, |\mathtt{r}|) = \frac{\pi^{d/2}}{(d+1)\cdot\Gamma(\frac{d}{2}+1)} \cdot \frac{(t^2 - |\mathtt{r}|^2 \cot^2(\theta))^{(d+1)/2}}{2^d} \cdot \tan^d(\theta) \tag{2.21}$$

We note that particularly for $\theta = \pi/4$, this reduces to a dimension dependent scaling of the metric interval in Minkowski spacetime.

Coming back to the problem of finding the probability of a link, we can say that

$$P(a \to x) = \exp\left(-\rho_c \cdot \mathrm{Vol}(t, |\mathtt{r}|)\right) \tag{2.22}$$

which is significant for $\mathrm{Vol}_d(t, |\mathtt{r}|) < V_c$, where V_c is the discretization scale. The region where this is significant is the light cone minus a hyperboloid of radius V_c. Upto height T in the time axis, the volume of this region is

$$\mathrm{Vol.(link)} = \frac{\pi^{d/2}}{(d+1)\cdot\Gamma(\frac{d}{2}+1)} \cdot T^{d+1} - \mathrm{Vol.(hyperboloid)} \tag{2.23}$$

where the volume of the hyperboloid is defined by the inequality

$$t^2 - \mathtt{r}^2 \geq V_c^2, \quad \text{with} \quad t \in [V_c, T],$$

where $\mathtt{r} = \sqrt{x_1^2 + \cdots + x_d^2}$ denotes the spatial distance from the origin. This defines a hyperboloid symmetric about the time axis t, and we are interested in computing the spacetime volume enclosed by this region up to height $t = T$.

At each fixed time $t > V_c$, the inequality implies

$$\mathtt{r}^2 \leq t^2 - V_c^2,$$

so the spatial slice is a d-dimensional ball of radius $R(t) = \sqrt{t^2 - V_c^2}$. The volume of this ball is

$$\mathrm{Vol}_d(R(t)) = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)} \left(t^2 - V_c^2\right)^{d/2}.$$

To compute the total spacetime volume, we integrate the spatial volume over time from V_c to T:

$$\mathrm{Vol}_{d+1}(V_c, T) = \int_{V_c}^{T} \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)} \left(t^2 - V_c^2\right)^{d/2} dt. \tag{2.24}$$

We can simplify this integral via the substitution $t = V_c \cosh u$, so that $t^2 - V_c^2 = V_c^2 \sinh^2 u$

and $dt = V_c \sinh u \, du$. Then the volume becomes

$$\text{Vol(hyperboloid)} = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)} \int_0^{\cosh^{-1}(T/V_c)} (V_c \sinh u)^{d+1} \, du \tag{2.25}$$

$$= \frac{\pi^{d/2} V_c^{d+1}}{\Gamma\left(\frac{d}{2}+1\right)} \int_0^{\cosh^{-1}(T/V_c)} \sinh^{d+1} u \, du. \tag{2.26}$$

By taking the derivative of Vol(link) with respect to T, we see that it is a monotonically increasing function. Thus, the possible region for a link to exist is infinite; we will surely find a link in the infinite future. Since links are nearest neighbors of each other, this implies that there are infinitely many links in the far far future. This gives rise to an emergent non-locality in the causal set.

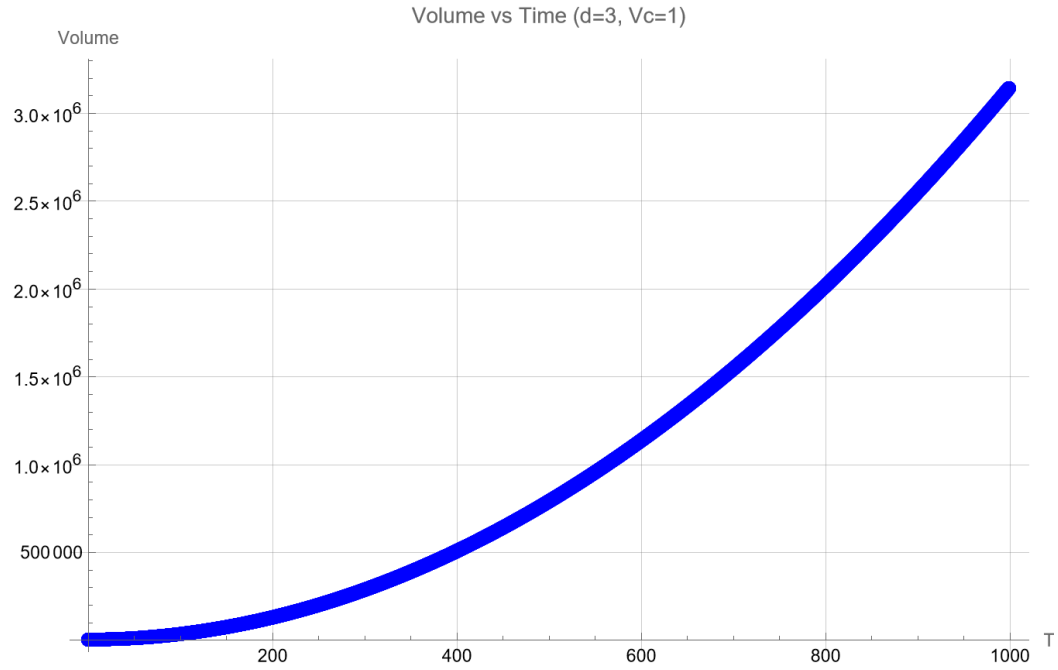


Figure 2.2: This is the volume of between the light cone and the hyperboloid of radius V_c in a 3D Minkowski spacetime. It tells us that there are infinitely many links in the far future.

2.2.3 Invariance in Random Sampling

Randomly and uniformly sampling a set of point from a Euclidean manifold preserves the continuous symmetries because the action of the symmetry group results in non-trivial orbits which provide a fibration of the ensemble of Poisson sprinklings. On average, the expected properties behave like properties of the underlying manifold. This argument rests on the fact that the symmetry group of a Euclidean manifold is compact.

A given realization of the Poisson process breaks the continuous symmetries of the manifold. To show this, consider the concrete example of a plane $\mathbb{R}^2$ equipped with a Euclidean metric. Let $p : \mathcal{C}(\mathbb{R}^2, \rho) \to \mathbb{R}^2$ be the natural embedding map, where $\mathcal{C}(\mathbb{R}^2, \rho)$ is the ensemble of Poisson

sprinklings of points in $\mathbb{R}^2$ at density ρ.

Rotations $r \in SO(2)$ act on the points in $\mathbb{R}^2$ and translations $t \in \mathbb{R}^2$ induce maps $r^* = p^{-1} \circ r \circ p$ and $t^* = p^{-1} \circ t \circ p$ on the sprinkling ensemble. For $P \in \mathcal{C}(\mathbb{R}^2, \rho)$, choose some element e in $p(P)$. The set of unit vectors in $\mathbb{R}^2$ is S^1 (which is compact). An element $\mathbf{d}$ of S^1 transforms as $\mathbf{d} \mapsto r^*(\mathbf{d})$ under some rotation $r \in SO(2)$.

We want a rule to assign directions to each P in $\mathcal{C}(\mathbb{R}^2, \rho)$. A natural choice is to fix $\mathbf{d}$ to be the direction of the vector from e to its closest neighbor in $p(P)$. Because the probability of finding two points at exactly the same radius is zero in a Poisson process, this is well defined. We can define a *direction map* $\mathbf{D}_e : \mathcal{C}(\mathbb{R}^2, \rho) \to S^1$ for a fixed point e in $p(P)$ such that this map commutes with any rotation $r \in SO(2)$.

Associated with $\mathcal{C}(\mathbb{R}^2, \rho)$ is a probability distribution μ arising from the Poisson sprinkling which assigns to every measurable set A in $\mathcal{C}(\mathbb{R}^2, \rho)$ a probability $\mu(A) \in [0, 1]$. Since the Poisson process is volume preserving, the measure is invariant under the action of the symmetry group, i.e. $\mu = \mu \circ r$. We can now use the direction map to define a measure on the unit circle S^1 as $\mu_{\mathbf{D}} = \mu \circ \mathbf{D}^{-1}$. Now, this measure is invariant under the action of symmetry group because of the invariance of μ and the commutativity of the direction map with the action of the symmetry group. Because S^1 is compact, we can choose n non-overlapping measurable sets of equal size, say $A_1, A_2, \ldots, A_n$, which results in $\mu_{\mathbf{D}}(A_i) = \frac{1}{n}$ for all i. Hence, it is possible to consistently assign a direction in a Euclidean manifold and break the continuous symmetries of the manifold in a Poisson process.

In a Lorentzian manifold, however, the symmetry group is non-compact, due to boosts, and the directions lie on a hyperboloid instead of a sphere (or an appropriate compact generalization thereof). Let's assume that a consisten direction map does exist. Now, we cannot cover the hyperboloid with a finite number of non-overlapping measurable sets of equal size; if we could we would have $\mu_{\mathbf{D}}(A_i) = 0$ for all i and the sum to be one. This is a contradiction. Hence, we cannot consistently assign a direction in a Lorentzian manifold. This means that the continuous symmetries of a Lorentzian manifold are not broken by the Poisson process. *Do distributions other than the Poisson process also preserve the volume?* If they do, they would also predict local Lorentz Invariance. We could open ourselves to a richer variety of statistics if we try to find distributions that preserve volume. The most suitable from such distributions would be the ones that minimise fluctuations.

We end this section by noting that three reasons why every realization of a Poisson process on a Lorentzian manifold preserves the continuous symmetries because: (i) the symmetry group is non-compact, (ii) the directions lie on a hyperboloid which is non-compact, and (iii) the Poisson sampling preserves the volume measure.

2.3 What Are We Looking For?

Causal set theory proposes that spacetime is fundamentally discrete rather than continuous - made up of individual events connected by causal relationships, like a cosmic network where some events can influence others. This approach to quantum gravity suggests that the smooth spacetime we observe emerges from an underlying discrete structure, much like how a computer image appears smooth but is actually made of individual pixels.

2.3.1 Can We Reconstruct Spacetime from Causal Structure?

The first major question in causal set theory is whether we can work backwards: if we know which events can causally influence which other events, can we figure out where and when everything happened in spacetime?

The researchers developed two different methods for measuring 'distances' in this causal network. One approach counts the longest chains of causal connections between events, similar to finding the longest path through a social network. The other method looks at the 'volume' of causal regions - essentially measuring how much spacetime sits between two events. To make this reconstruction unique, they had to break some symmetries. They fixed certain reference points in their spacetime region, like setting coordinates on a map by designating specific landmarks. This allowed them to assign actual spacetime coordinates to events based purely on their causal relationships.

The remarkable result is that both methods work. You can indeed reconstruct spacetime geometry from causal structure alone, though with some inevitable uncertainty. One method gives better accuracy in coordinate placement, while the other preserves important structural properties better. This proves that causal relationships contain nearly all the geometric information we need to describe spacetime. One method that I did not discuss, and which can be developed from my presented techniques, is to assign time coordinates based on either geodesic distance or Alexandrov intervals, and to assign space coordinates based on the antichain property analogous to the time coordinate assignment. Interestingly, this method is the best of the three, as shown in a paper.

2.3.2 What About Large Causal Sets?

When researchers analyzed small causal sets with just a few events, they discovered both encouraging and troubling patterns. The good news was that their theoretical predictions about how these discrete spacetimes should behave were confirmed by detailed calculations.

However, a serious problem emerged. When you count all possible ways to arrange causal relationships among events, most arrangements look nothing like the smooth spacetime we observe. Instead, they tend to form layered structures - like having most events clustered at a few

distinct times with big gaps in between. This creates an *entropy crisis* because the number of possible arrangements grows much faster than it should for a realistic physical system.

To solve this, researchers proposed giving different causal arrangements different probabilities, similar to how some configurations in statistical mechanics are more likely than others. By penalizing arrangements that deviate too much from smooth spacetime, they hoped to favor more realistic configurations. Early attempts at this approach showed promise but left many questions unanswered about how to choose the right weighting scheme.

2.3.3 Why Does Causal Structure Matter So Much?

The mathematical foundations reveal why the causal set approach makes sense. In Einstein's general relativity, nearly all of spacetime's geometric properties are encoded in its causal structure - which events can influence which others. Several powerful theorems show that if you preserve causal relationships, you automatically preserve almost all geometric relationships too.

This means causality isn't just one aspect of spacetime – it's the fundamental organizing principle. The causal network between events determines distances, angles, curvature, and almost everything else we care about geometrically. Only one small piece of information – the overall *volume scale* – isn't captured by pure causal structure, and this might be determined by the discrete nature of causal sets.

This mathematical insight provides strong justification for the causal set approach. If causal structure contains nine-tenths of spacetime's geometric information, then starting with causal relationships and letting geometry emerge makes perfect sense.

2.3.4 How to Distribute Events in Spacetime?

The final piece involves understanding how discrete events should be distributed throughout spacetime. Rather than placing events in regular patterns (which would break the symmetries of relativity), researchers use random Poisson processes, randomly scattering points while maintaining uniform density on average.

This random approach automatically preserves the symmetries of relativity, which is crucial for physical realism. Unlike in ordinary space where you can consistently choose preferred directions, in spacetime the non-compact nature of relativistic symmetries makes this impossible. The random approach naturally handles this difficulty.

However, this randomness creates new problems. In infinite spacetime, random processes guarantee that arbitrarily large empty regions will occur somewhere, creating a *void problem*. Even worse, the probability calculations predict that distant events will eventually become causally connected, violating the locality principle that nearby events should only directly influence nearby events.

2.3.5 Current State and Future Challenges

These four investigations establish causal set theory as a mathematically sophisticated approach to quantum gravity with both remarkable successes and significant challenges. The theory successfully shows how spacetime geometry can emerge from discrete causal structure while preserving relativistic symmetries. The coordinate reconstruction methods work, confirming that causal relationships really do contain geometric information.

Yet fundamental problems remain. The entropy crisis suggests that most random discrete spacetimes look nothing like smooth space, requiring sophisticated statistical mechanics to explain why we observe smooth geometry. The non-locality problem threatens the basic principle that influences should propagate locally through space and time. The void problem raises questions about whether the universe must be finite or whether we live in an anomalously smooth region.

Despite these challenges, the causal set approach provides a promising framework for understanding quantum gravity. It emphasizes the central role of causality in spacetime structure and offers a mathematically rigorous way to explore how discrete events can give rise to the smooth geometry we observe.

Chapter 3

Enumeration of Posets

We now discuss in detail how to count the number of posets with a given number of elements. This is a very important problem in CST because we will soon see that the entropy of non-manifold like posets cannot be ignored if we want to construct the path integral on the space of all posets. We begin with an analysis of Kleitman and Rothschild's work on the number of finite topologies and show its relation to the counting of posets. Dhar's work on the construction of entropy on these structres and the resulting first order transitions follows. We end this chapter with a modification of Dhar's entropy construction to include higher-order effects and numerical results are presented.

3.1 Equivalent Enumeration Problems

All logarithms in this section are to base 2.

Consider the following sets; X is a set of n elements, $T(X)$ is the set of all distinct topologies on X, $P(X)$ is the set of distinct partial orders on X, $T_0(X)$ is the set of all distinct T_0-topologies on X (i.e. if a, b are elements of X, then there is some open set containing one but not both of them), $O(X)$ is the set of all preorders on X.

We denote $T_n = |T(X)|$, $P_n = |P(X)|$, $T_{n,0} = |T_0(X)|$, and $O_n = |O(X)|$. The number of isomorphism classes are denoted as T'_n, P'_n, $T'_{n,0}$, and O'_n.

Lemma 9. *$T_n = O_n$, $T_{n,0} = P_n$, $T'_n = O'_n$, and $T'_{n,0} = P'_n$. All eight of these quantities have logarithms (to base 2) that are asymptotically equal as n tends to infinity.*

Proof. We use the method used in 'The number of topologies on n points' by S. D. Chatterji. To define a topology on the set X, we use the method of closure operators satisfying Kuratowski's postulates. Since X is finite, the closure operator is uniquely and completel determined by the functions F from X to the $P(X)$, the power set of X, such that:

1. $\forall a \in X$, $a \in F(a)$;

2. if $y \in F(x)$, then $F(y) \subset F(x)$.

If we further define $F(A) = \bigcup\{F(a) : a \in A\}$, for any $A \subseteq X$, then F follows the closure postulates; $F(\emptyset) = \emptyset$, $A \subseteq F(A)$, $F(F(A)) = F(A)$, and $F(A \cup B) = F(A) \cup F(B)$. Our problem reduces to counting the number of such functions F from X to $P(X)$.

The function F induces a T_0 topology iff given any two distinct points $a, b \in X$, either $a \notin F(b)$ or $b \notin F(a)$ or both. Let's now define a relation $<$ such that $x < y$ if $x \in F(y)$. This is verified to be relfexive and transitive, i.e. it is a preorder. For any given preorder R, we can define a closure operator F_R such that $F_R(x) = \{y \in X : xRy\}$. Hence, we have that $T_n = O_n$.

If we demand the relation to be antisymmetric as well, then the function induces a T_0 topology and the relation is a partial order. This means that $T_{n,0} = P_n$. Other two equalities follow trivially.

To prove the second part of the lemma, we note that to obtain a preorder on a set X with n elements we partition X into subsets (in at most $n!2^{n-1}$), and then partially order the subsets (in at most P_n ways). Hence, $\log(O_n/P_n) = \mathcal{O}(n^2)$.

Then, we note that there are at most $n!T'_n$ distinct topologies for T'_n isomorphism classes, because we can permute the elements of X in $n!$ ways, and similarly for $T'_{n,0}$. Hence, we have $\log(T_n/T'_n) = \mathcal{O}(n^2)$ and $\log\left(T_{n,0}/T'_{n,0}\right) = \mathcal{O}(n^2)$.

Finally, we note that minimum number of P_n distinct partial orders on a set of n elements is $2^{n^2/4}$ because we can construct two layers, as in chapter 1, and make a bipartite graph with edges directed from one layer to the other. For more layers, we can have more of these and hence, we have $P_n \geq 2^{n^2/4}$. □

Theorem 3. *For some constant C, we have $\log P_n \leq n^2/4 + Cn^{3/2} \log n$ for all n.*

This theorem is important because we now have both upper and lower bounds on the number of posets. The proof is non-trivial and instructive. This can be proved by induction. We begin by constructing four types of (Hasse) diagrams of posets which will be used in the proof. For Q a set of vertices in a diagram, define $N(Q)$ to be the set of vertices in the diagram that are adjacent to any vertex in Q. We say a vertex is *minimal* if does not cover any other vertex in the diagram.

A For a set V, let $A(V)$ be the set of diagrams of $|V|$ vertices such that each element of V is adjacent to at most $(|V| - 1)/64$ other elements of V. When $|V| = k$, we denote $A_k = |A(V)|$.

B For a set V, let $B(V)$ be the set of diagrams of $|V|$ vertices with a minimal element covered by a set Q of size $\lfloor(|V| - 1)^{1/2}\rfloor$, with $N(Q)$ containing at least $(|V| - 1)/2$ vertices. When $|V| = k$, we denote $B_k = |B(V)|$.

C For a set V, let $C(V)$ be the set of diagrams of $|V|$ vertices with a minimal element covered by a set Q of size $\lfloor(|V|-1)^{1/2}\rfloor$, with $N(Q)$ containing at most $(|V|-1)/2$ vertices. When $|V|=k$, we denote $C_k = |C(V)|$.

Lemma 10. $\log(A_{n+1}/P_n) < n/5$ *for sufficiently large* n.

Proof. We begin with some n vertex diagram and add a vertex such that at most $n/64$ vertices are covered by it and such that this vertex is covered by at most $n/64$ vertices. The number of ways to choose a diagram on the n points is $\binom{n+1}{n}P_n$ and we can choose the points that are covered by the new vertex

$$\sum_{i=0}^{\lfloor n/64\rfloor}\binom{n}{i} \leq \frac{n}{64}\cdot\binom{n}{\lfloor n/64\rfloor} \tag{3.1}$$

Hence, we have

$$\log(A_{n+1}/P_n) \leq \log(n+1) + 2\log\frac{n}{64} + 2\log\binom{n}{\lfloor n/64\rfloor} \tag{3.2}$$

Using Stirling's approximation, we can simplify this to the lemma. □

Lemma 11. $\log(B_{n+1}/P_n) < n/2 + n^{1/2}\log n$ *for sufficiently large* n.

Proof. In this case, we cannot form any diagram with triangles regardless of the orientation of edges. Choose a configuration of n vertices in $\binom{n+1}{n}P_n$ ways. Add a minimal vertex which is covered by a set Q of size $\lfloor n^{1/2}\rfloor$ vertices, such that $N(Q) \geq n/2$. There are $n/2 - \lfloor n^{1/2}\rfloor$ vertices that are not in either Q or $N(Q)$. We can choose to either connect or not connect these vertices to the new vertex. Thus, in total, we have

$$\log(B_{n+1}/P_n) \leq \log(n+1) + \log\binom{n}{\lfloor n^{1/2}\rfloor} + \frac{n}{2} - \lfloor n^{1/2}\rfloor \tag{3.3}$$

Choosing the largest terms, and using Stirling's approximation, we have the lemma. □

Lemma 12. $\log\Big(C_{n+1}/P_{n-\lfloor n^{1/2}\rfloor}\Big) \leq n^{3/2}/2 + 4n$ *for sufficiently large* n.

Proof. This proof is not so trivial. Choose $|V| = n+1$. Then, choose a diagram H on $n - \lfloor n^{1/2}\rfloor$ vertices of V. Following this, choose one element of the remaining to be minimal and the rest to form the set Q. Now, cover the minimal element by the set Q, and connect Q to at most $n/2$ vertices in H such that no forbidden configuration occurs (triangles and quadrilaterals in this case). Finally, cover the minimal element with some of the elements in $N(Q)$.

This leads to the following inequality

$$\log\Big(C_{n+1}/P_{n-\lfloor n^{1/2}\rfloor}\Big) < n + 1 + \log n + \frac{n^3/2}{2} + \frac{n}{2} + \log n + n + n \tag{3.4}$$

There is subtle counting involved in this and I refer the reader to Kleitman and Rothschild's paper for the details. □

Having established these lemmas, the induction is now straightforward. We need to note that all possible diagrams lie in these classes for sufficiently large n. In this section, we have established bounds on the number of posets possible in a set of n elements. We see that the entropy of the possible number of posets is at least quadratic in n. From the theory of Poisson processes, we know that the entropy should be linear in n. This indicates that the kinematics of causal sets (posets) is much more complicated than what meets the eye. To understand this better, we need to formalise the notion of entropy in posets.

3.2 Entropy With Constraints - I

Right now, we know that for a set with n elements, the number of distinct partial orders lies in

$$n^2/4 \leq \log_2 P_n \leq n^2/4 + Cn^{3/2} \log n \tag{3.5}$$

for some finite constant C. Defining the entropy of this set as

$$S = \lim_{n\to\infty} \frac{2}{n^2} \ln P_n \tag{3.6}$$

we obtain that $S = \frac{1}{2} \ln 2$. One must note that this is maximum of the entropy associated with with Bernoulli distribution.

Let's take this model one step further. Suppose we know the fraction of comparable pairs in a poset with n elements to be ρ. This means that the number of comparable pairs is $\rho n^2/2$ in the large n limit. Define an entropy function

$$S(\rho) = \lim_{n\to\infty} \frac{2}{n^2} \ln P_n(\rho) \tag{3.7}$$

We know that $S(0) = 0$ and $S(1) = 0$. Since the entropy is definitely reduced if we specify the fraction of comparable pairs, we have $S(\rho) \leq \frac{1}{2} \ln 2$. An important question is *does* $S(\rho)$ *ever reach its maximum?* The analysis of this question is done in Dhar's paper and is quite interesting.

3.2.1 General Features of the Entropy Function

Lemma 13. *$S(\rho)/\rho$ is a monotonic non-increasing function of ρ.*

Proof. Take a poset with n elements and ρ fraction of comparable pairs. Add ϵ new elements that are incomparable to all the existing elements and to each other. In the layer picture, it

would look like

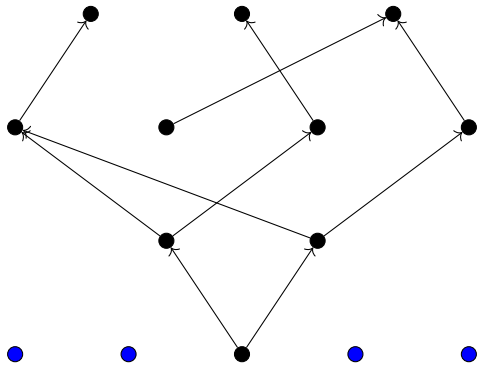


Figure 3.1: A layered representation of a poset, with the new elements (in blue) added at the bottom layer. The arrows indicate the direction of the partial order.

The new density of comparable pairs is $\rho' = \frac{\rho n(n-1)}{(n+\epsilon)(n+\epsilon-1)}$. Thus we have trivially that

$$P_n(\rho) \leq P_{n+\epsilon}(\rho') \tag{3.8}$$

If we take the logarithm of both sides and take the limit of both n and ϵ to infinity, we get

$$\frac{S(\rho)}{\rho} \leq \frac{S(\rho')}{\rho'} \tag{3.9}$$

for $\rho' < \rho$. □

Lemma 14. *$S(\rho)/(1-\rho)$ is a monotonic non-decreasing function of ρ.*

Proof. Instead of adding ϵ new elements that are incomparable to all the existing elements and to each other, we add ϵ new elements that are comparable to all the existing elements and to each other, that is, add a chain below a minimal element.

$$\rho' = \frac{2}{(n+\epsilon)(n+\epsilon-1)}\left(\rho\frac{n(n-1)}{2} + n\epsilon + \frac{\epsilon(\epsilon-1)}{2}\right) \tag{3.10}$$

Now, we have $P_n(\rho) \leq P_{n+\epsilon}(\rho')$ again. Taking the logarithm of both sides, and taking the limit of both n and ϵ to infinity, combined with $\rho' > \rho$, proves the lemma. □

Lemma 15. *Any function $S(\rho)$ that satisfies the above two lemmas is a continuous function of ρ.*

Proof. We divide the proof into two parts: continuity on the open interval $(0, 1)$, and continuity at the endpoints $\rho = 0$ and $\rho = 1$.

These lemmas provide upper and lower bounds on the difference quotient of $S(\rho)$.

Let $0 < \rho_1 < \rho_2 < 1$. Then, from the first lemma, we have:

$$\frac{S(\rho_2) - S(\rho_1)}{\rho_2 - \rho_1} \leq \frac{S(\rho_1)}{\rho_1} \tag{3.11}$$

and from the second lemma, we have:

$$\frac{S(\rho_2) - S(\rho_1)}{\rho_2 - \rho_1} \geq \frac{S(\rho_2)}{1 - \rho_2} - \frac{S(\rho_1)}{1 - \rho_1} \tag{3.12}$$

This shows that the difference quotient of S is bounded above and below by finite expressions depending on ρ_1, ρ_2. Hence, $S(\rho)$ is locally Lipschitz on $(0, 1)$, and therefore continuous on $(0, 1)$. Since the limits on $\rho = 0$ and $\rho = 1$ are well-defined, and extend, we can conclude that $S(\rho)$ is continuous on the closed interval $[0, 1]$. □

Alternate Proof. We aim to show that for every $\rho \in [0, 1]$, $S(\rho)$ is continuous, i.e.,

$$\forall \epsilon > 0,\ \exists\ \delta > 0 \text{ such that } |\rho' - \rho| < \delta\ \Rightarrow\ |S(\rho') - S(\rho)| < \epsilon. \tag{3.13}$$

Let us assume $\rho \in (0, 1)$, and consider a small increment $\epsilon > 0$ such that $\rho + \epsilon < 1$. From the monotonicity of $\frac{S(\rho)}{\rho}$ and $\frac{S(\rho)}{1-\rho}$, we know:

$$\frac{S(\rho + \epsilon)}{\rho + \epsilon} \leq \frac{S(\rho)}{\rho}, \qquad \frac{S(\rho + \epsilon)}{1 - (\rho + \epsilon)} \geq \frac{S(\rho)}{1 - \rho} \tag{3.14}$$

Multiplying the first inequality by $\rho(\rho + \epsilon)$ gives:

$$\rho \cdot S(\rho + \epsilon) \leq (\rho + \epsilon) \cdot S(\rho) \quad \Rightarrow \quad S(\rho + \epsilon) \leq \frac{\rho + \epsilon}{\rho} S(\rho) \tag{3.15}$$

Similarly, multiplying the second inequality by $(1 - \rho)(1 - \rho - \epsilon)$ gives:

$$(1 - \rho) \cdot S(\rho + \epsilon) \geq (1 - \rho - \epsilon) \cdot S(\rho) \quad \Rightarrow \quad S(\rho + \epsilon) \geq \frac{1 - \rho - \epsilon}{1 - \rho} S(\rho) \tag{3.16}$$

So we have:

$$\frac{1 - \rho - \epsilon}{1 - \rho} S(\rho) \leq S(\rho + \epsilon) \leq \frac{\rho + \epsilon}{\rho} S(\rho) \tag{3.17}$$

Now, subtracting $S(\rho)$ from all parts:

$$\left(\frac{1 - \rho - \epsilon}{1 - \rho} - 1 \right) S(\rho) \leq S(\rho + \epsilon) - S(\rho) \leq \left(\frac{\rho + \epsilon}{\rho} - 1 \right) S(\rho) \tag{3.18}$$

Simplifying both sides:

$$-\frac{\epsilon}{1-\rho}S(\rho) \leq S(\rho+\epsilon) - S(\rho) \leq \frac{\epsilon}{\rho}S(\rho) \tag{3.19}$$

Therefore,

$$|S(\rho+\epsilon) - S(\rho)| \leq \epsilon \cdot \max\left(\frac{1}{\rho}, \frac{1}{1-\rho}\right) S(\rho) \tag{3.20}$$

This gives a uniform bound on how much $S(\rho+\epsilon)$ can deviate from $S(\rho)$, in terms of ϵ. Since $S(\rho)$ is bounded (by $\frac{1}{2}\ln 2$), we can write:

$$|S(\rho+\epsilon) - S(\rho)| \leq C\epsilon \quad \text{for some constant } C = \max\left(\frac{1}{\rho}, \frac{1}{1-\rho}\right) S(\rho) \tag{3.21}$$

Hence, given any $\varepsilon > 0$, choose

$$\delta = \frac{\varepsilon}{C} \Rightarrow \text{if } |\rho' - \rho| < \delta, \text{ then } |S(\rho') - S(\rho)| < \varepsilon \tag{3.22}$$

This proves that $S(\rho)$ is continuous at every $\rho \in (0,1)$. The boundary points $\rho = 0$ and $\rho = 1$ can be handled exactly as in the first proof, since we know:

$$\lim_{\rho\to 0^+} S(\rho) = 0 = S(0), \qquad \lim_{\rho\to 1^-} S(\rho) = 0 = S(1) \tag{3.23}$$

Thus, $S(\rho)$ is continuous on all of $[0,1]$. □

Conjecture 1. *The entropy function $S(\rho)$ is a convex function of ρ.*

Proof. This is not a complete proof, but a try towards one. Convexity is defined as a function f being convex if for all x_1, x_2 in the domain and $\lambda \in [0,1]$, we have:

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2) \tag{3.24}$$

Consider a poset with n elements and a fraction ρ of comparable pairs. Let there be a number $\lambda \in [0,1]$. We divide our poset into two parts: one with $\lambda n = n_1$ elements and the other with $(1-\lambda)n = n_2$ elements. Poset 1 has a fraction ρ_1 of comparable pairs and poset 2 has a fraction ρ_2 of comparable pairs. We know that the total number of comparable pairs in the combined poset is given by $\rho\binom{n}{2}$ and the two parts contribute $\rho_1\binom{n_1}{2} + \rho_2\binom{n_2}{2}$. In the large n limit, we can write this as:

$$\text{Number of comparable pairs} = \rho\frac{n^2}{2} \tag{3.25}$$

$$= \left(\lambda^2\rho_1 + (1-\lambda)^2\rho_2\right)\frac{n^2}{2} + \text{cross terms} \tag{3.26}$$

Hence, we know $\rho \geq \lambda^2\rho_1 + (1-\lambda)^2\rho_2$. I would like to think that this may help somehow, along with Lemma 1, to show that $S(\rho)$ is convex. I am not sure of the method following this. □

3.2.2 Nature of the Solution

We now build this from a variational principle. Consider r number of layers, with the i-th layer containing f_i fraction of elements. Between each adjacent layer, we choose α fraction of pairs to be comparable. Any two layers separted by more than one layer automatically have all elements comparable. The set $\{f_i\}$ is a partition of unity. Counting the number of incomparable pairs in each layer, and the number of incomparable pairs between adjacent layers,

$$(1-\rho)\binom{n}{2} = \sum_{i=1}^{r}\binom{f_i n}{2} + \sum_{i=1}^{r-1}(1-\alpha)f_i f_{i+1}n^2 \tag{3.27}$$

$$1-\rho = \sum_{i=1}^{r} f_i^2 + 2\sum_{i=1}^{r-1}(1-\alpha)f_i f_{i+1} \tag{3.28}$$

For ease of notation, I will call $\sum_{i=1}^{r} f_i^2 = A$ and $\sum_{i=1}^{r-1} f_i f_{i+1} = B$, and for later use, I will replace $1-\rho$ with ρ_0 (the density of incomparable pairs). We can rewrite the above equation as:

$$\rho_0 = A + 2(1-\alpha)B \tag{3.29}$$

The total number of comparable pairs at density ρ is greater than or equal to $\binom{n^2B}{\alpha B}$ because of our demand that non-adjacent layers are completely comparable. Hence, the entropy has the following relation

$$S(\rho_0) \geq \lim_{n\to\infty}\frac{2}{n^2}\ln\binom{n^2B}{\alpha B} \tag{3.30}$$

We now use Stirling's approximation to simplify this expression.

$$S(\rho_0) \geq 2B\cdot\mathcal{S}(\alpha) \tag{3.31}$$

where $\mathcal{S}(\alpha) = -\alpha\ln\alpha - (1-\alpha)\ln(1-\alpha)$ is the entropy of a Bernoulli distribution with parameter α.

All we know about this entropy function is that it is a continuous function of ρ_0, that $S(0) = 0$, $S(1) = 0$, that $S(\rho_0)/\rho_0$ is a monotonic non-decreasing function of ρ_0, and that $S(\rho_0)/(1-\rho_0)$ is a monotonic non-increasing function of ρ_0, and that

$$\frac{1}{2}\ln 2 \geq S(\rho_0) \geq 2B\cdot\mathcal{S}(\alpha) \tag{3.32}$$

It is now imperative to note that since $\{f_i\}$ is a partition of unity, we get that B is maximized at $1/4$ and minimized at 0. This can be obtained in two different ways; (a) two layers and both

containing 1/2 fraction of elements, (b) three layers with 1/2 fraction in the middle layer and the sum of the other two layers being 1/2. If we choose α to be 1/2 in addition to either of these cases, we will have the maximum entropy. This condition is particularly interesting because we can never, by construction, have a layered poset with $\mathcal{S}(\alpha) > \ln 2$ or with $B > 1/4$, and this highly constrains the types of posets with maximum entropy. If we have a finite section of ρ_0 for which we have maximum entropy (only possible with a 3-layer poset) then we expect a *phase transition* to occur at the end points of this interval, before and after which this function is not a horizontal line.

Lemma 16. *$S(\rho_0) = \frac{1}{2}\ln 2$ for $\rho_0 \in [5/8, 3/4]$.*

Proof. We can show this trivially by choosing a 3-layer poset ($\alpha = 1/2$) with 1/2 fraction of elements in the middle layer and with

$$f_1 = \frac{1}{4} - \Delta(\rho_0) \quad \text{and} \quad f_3 = \frac{1}{4} + \Delta(\rho_0) \tag{3.33}$$

such that

$$\rho_0 = \frac{1}{4} + 2\left(\frac{1}{16} + \Delta(\rho_0)^2\right) + 2 \times \frac{1}{2} \times \frac{1}{4} \tag{3.34}$$

which results in

$$\Delta(\rho_0) = \left(\frac{\rho_0}{2} - \frac{5}{16}\right)^{1/2} \tag{3.35}$$

The requirement than $f_1 > 0$ and that the quantity inside the square root is non-negative gives us the bounds on ρ_0 as $5/8 \leq \rho_0 < 3/4$. $\rho_0 = 3/4$ is the case when we have a 2-layer poset with 1/2 fraction of elements in each layer. □

Now, since we have found a finite section of ρ_0 for which the entropy is maximum, we can conclude that there is a phase transition at the end points of this section! Let's now try to understand the nature of $S(\rho_0)$ outside this section.

Lemma 17. *For $\rho_0 < 5/8$, $S(\rho_0) \geq \frac{4}{5}\rho_0 \ln 2$.*

Proof. This is shown easily using the monotonicity of $S(\rho_0)/\rho_0$. □

Lemma 18. *For $\rho_0 > 3/4$, $S(\rho_0) = \frac{1}{2}\mathcal{S}(\alpha)$ with $\alpha = 2(1 - \rho_0)$.*

Proof. Consider a two layer poset with 1/2 fraction of elements in each layer. Then, $\rho_0 = 1/2 + (1 - \alpha)/2$. □

We can go further to show some more properties of this function. To do this, let's construct a Lagrange Multiplies system of equations. We want to maximize the lower bound and find the

corresponding $\{f_i, \alpha\}$.

$$\mathcal{L}(\{f_i\}, \alpha, \lambda_1, \lambda_2) = 2B \cdot \mathcal{S}(\alpha) + \lambda_1 \left(1 - \sum_{i=1}^{r} f_i\right) + \lambda_2 \left(\rho_0 - A - 2(1-\alpha)B\right) \tag{3.36}$$

which leads to

$$\frac{\partial \mathcal{L}}{\partial \alpha} = 2B \ln\left(\frac{1-\alpha}{\alpha}\right) + 2\lambda_2 B \tag{3.37}$$

$$\frac{\partial \mathcal{L}}{\partial f_i} = 2\left(\mathcal{S}(\alpha) - \lambda_2(1-\alpha)\right)\left(f_{i-1} + f_{i+1}\right) - \lambda_1 - 2\lambda_2 f_i \tag{3.38}$$

and the constraint equations, along with the obvious constraints $f_0 = f_{r+1} = 0$.

$$\alpha = \frac{1}{1 + e^{-\lambda_2}} \tag{3.39}$$

is one thing for certain. We also have an equation of the form

$$f_{i-1} + f_{i+1} = \mathfrak{A} + \mathfrak{B} f_i \tag{3.40}$$

which has a general solution where, since $\mathfrak{B} < 2$, we have symmetry about the middle layer. This means that we can have a 3-layer poset with $f_1 = f_3$ and $f_2 = 1 - 2f_1$. We can say with certainity that when we change the number of layers from r to $r+1$, the maximal variational entropy is more in the $r+1$ layer for smaller values of ρ_0 and is more in the r-layer for larger values of ρ_0. The transition in number of layers is a phase transition in the entropy function. This is a first order phase transition, as the entropy is continuous at the transition point but its derivative is not. This can be verified for $r = 1, 2, 3$ but for higher r an analytic proof gets messy.

This result is of prime importance in the kinematics of CST. Density of comparable pairs drives phase transitions in the entropy of this system, which can be measured by the chemical potential. For causal sets of size N in d dimensions, we generally expect $\sim N^{1/d}$ number of layers (the average length of the maximal chain). In the large N limit, this too diverges and we expect the layered picture to wash out, being equivalent to a random sampling in this limit. We have so far formalised our idea of the first chapter in great detail.

When Dhar proposed this model, he suggested a lattice gas model where there are $\binom{n}{2}$ sites and each of them has three possible states; either be unoccupied, or have a relation in either direction. Due to the transitivity of a partial order, his idea results in a 3-body interaction of the lattice sites. He commented that "The transitions are governed by the strong, long range nature of the 3-body interaction here, not usually encountered in physical systems," but after dealing with causality and Lorentz invariance we can see that this sort of transition is fundamental in a theory of causal sets and spacetime because of the "non-locality" of a null geodesic.

Kleitman and Rothschild, following this, published an article where they showed that for

all ρ_0 greater than 3/4, two layers are sufficient to describe the poset. The general results are summarised in the figure below.

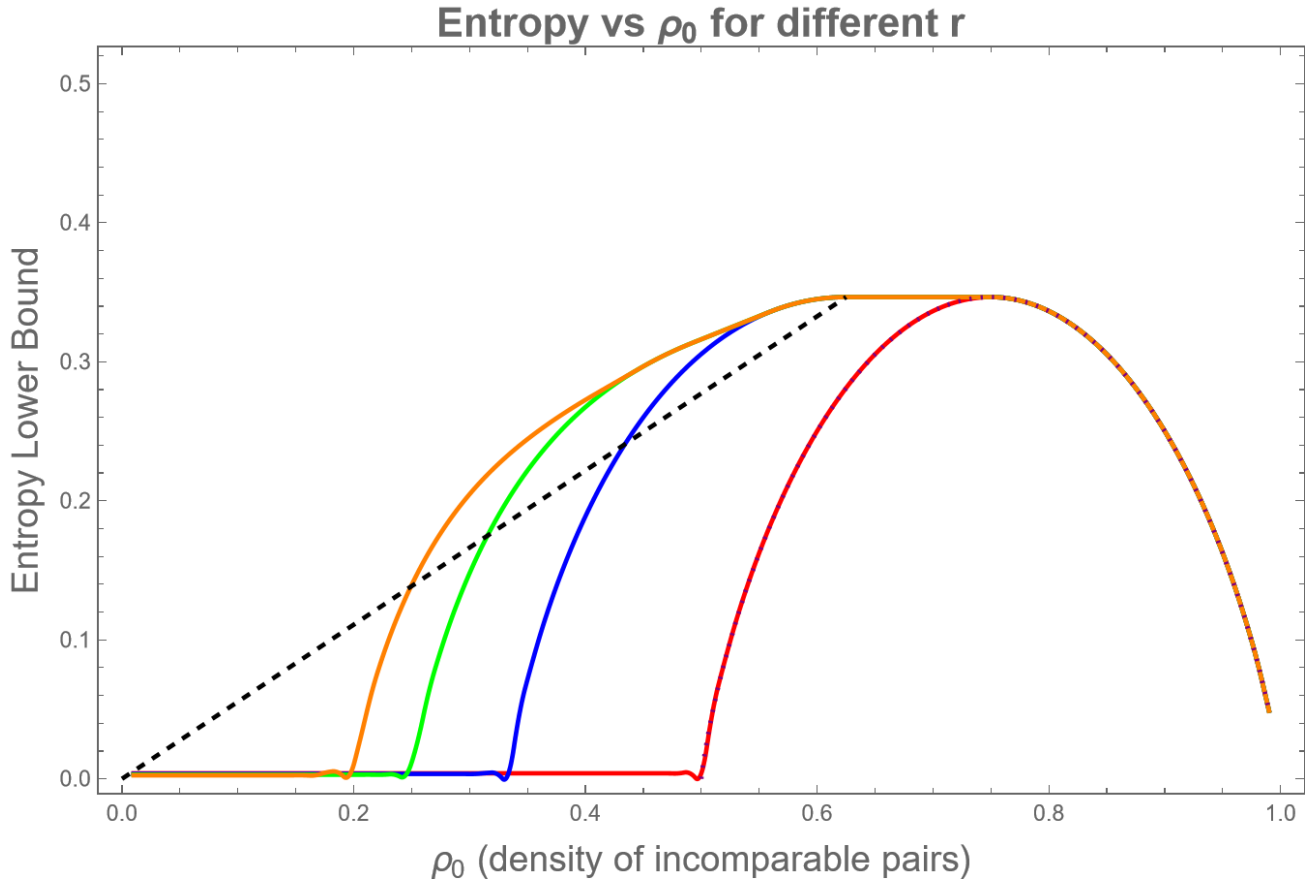


Figure 3.2: The variational entropy function $S(\rho_0)$ as a function of the density of incomparable pairs ρ_0, for different r. The envelope of these curves is the entropy function $S(\rho_0)$, which is continuous and has an infinite number of first order phase transitions. The dashed line is the lower bound upto $\rho_0 = 5/8$, and the dotted curve is the lower bound for $\rho_0 > 3/4$.

The graph above was obtained by plotting the maxima of the variational entropy function $S(\rho_0)$ for different number of layers r on Mathematica. We see the constant portion of the graph for $5/8 < \rho_0 < 3/4$. We also see that there may exist a $\rho_1 < 5/8$ till where the constant portion of the graph extends due to contributions of the higher layers. Right now, we do not have a definite value for ρ_1. The obtained graph has first order phase transitions where curves of different number of layers meet; the derivative jumps at these points.

3.3 Entropy With Constraints - II

Having established the qualitative nature of the phase transitions on the number of layers if we consider a variation on the density of incomparable pairs. We now try to generalize the idea by considering a variation on the density of links as well. The idea is to think of phase transitions in the entropy as a function of density of all possible chain lengths, zero, one, two, etc.

We obtain that $\rho_1 = 2\alpha B$ is the density of links in the poset, with α and B defined as before. We can try to constrain the region in the $\rho_0 - \rho_1$ plane where the entropy is well-defined, and try to identify the regions and curves corresponding to different number of layers.

3.3.1 General Features of the Solution

Just like we established the monotonicity relations with respect to ρ_0 in the previous section, we can note that for any fixed ρ_1, the same relations hold. We need to establish something similar for ρ_1 as well, keeping ρ_0 fixed.

Lemma 19. *$S(\rho_0, \rho_1)/\rho_1$ is a monotonic non-decreasing function of ρ_1 and $S(\rho_0, \rho_1)/(1 - \rho_1)$ is a monotonic non-increasing function of ρ_1 for fixed ρ_0.*

Proof. Consider a poset with n elements. We can add two layers of $nk/2$ elements each, where k is a positive scale, such that each pair in the new layers is comparable α fraction of the time.

We get that

$$\frac{\rho_0 n^2}{2}(1+k)^2 = \frac{\rho_0 n^2}{2} + \frac{n^2k^2}{4} + (1-\alpha)\frac{n^2k^2}{4} \tag{3.41}$$

$$\frac{\rho_1' n^2}{2}(1+k)^2 = \rho_1\frac{n^2}{2} + \alpha\frac{n^2k^2}{4} \tag{3.42}$$

This gives us that

$$\rho_1' - \rho_1 = \frac{(\rho_0+\rho_1)(1-(1+k)^2) + k^2}{(1+k)^2} \tag{3.43}$$

We know for sure that $P_n(\rho_0, \rho_1) \leq P_{n(1+k)}(\rho_0, \rho_1')$. Observe that the $\rho_0 + \rho_1 = 1$ case is handled by the previous section trivially. The remaining case of $\rho_0 + \rho_1 < 1$ is handled by choosing a k of our choosing such that $\rho_1' > \rho_1$ or $\rho_1' < \rho_1$ depending on the part of the lemma we are trying to prove. Note that the range of $\frac{k^2}{(1+k)^2-1}$ is $[0, 1)$. In the large n limit, we have now proved the lemma. □

Lemma 20. *$S(\rho_0, \rho_1)$ is a continuous function of ρ_0 and ρ_1.*

3.3.2 Formal Equations

We define the entropy function $S(\rho_0, \rho_1)$ via:

$$S(\rho_0, \rho_1) = \lim_{n\to\infty} \frac{2}{n^2} \ln P_n(\rho_0, \rho_1), \tag{3.44}$$

$$S_{\text{var}}(\rho_0, \rho_1) = 2B \cdot \mathcal{S}(\alpha), \tag{3.45}$$

where $\mathcal{S}(\alpha) = -\alpha \ln \alpha - (1-\alpha)\ln(1-\alpha)$ is the binary entropy function.

The parameters ρ_0 and ρ_1 are determined in terms of α and the layer weights f_i by:

$$\rho_0 = A + 2(1-\alpha)B, \tag{3.46}$$

$$\rho_1 = 2\alpha B, \tag{3.47}$$

where

$$A = \sum_{i=1}^{r} f_i^2, \tag{3.48}$$

$$B = \sum_{i=1}^{r-1} f_i f_{i+1}, \tag{3.49}$$

subject to the normalization and positivity constraints:

$$\sum_{i=1}^{r} f_i = 1, \tag{3.50}$$

$$f_i \geq 0, \quad \alpha \in (0,1). \tag{3.51}$$

We can then define a Lagrange multiplier system to maximize the entropy function:

$$\mathcal{L}(\{f_i\}, \alpha, \rho_0, \rho_1, \{\lambda_i\}) = S_{\text{var}}(\rho_0, \rho_1) + \lambda_1 \left(1 - \sum_{i=1}^{r} f_i\right) + \lambda_2 \left(\rho_0 - A - 2(1-\alpha)B\right) + \lambda_3 \left(\rho_1 - 2\alpha B\right) \tag{3.52}$$

We get the equations:

$$\frac{\partial \mathcal{L}}{\partial \alpha} = 0 \implies 2B\left(\ln\left(\frac{1-\alpha}{\alpha}\right) + \lambda_2 - \lambda_3\right) = 0, \tag{3.53}$$

$$\frac{\partial \mathcal{L}}{\partial f_i} = 0 \implies 2\left(\mathcal{S}(\alpha) - \lambda_2(1-\alpha) - \lambda_3\alpha\right)(f_{i-1} + f_{i+1}) - \lambda_1 - 2\lambda_2 f_i = 0. \tag{3.54}$$

We know for sure that $\rho_0 + \rho_1 > 1$ is a disallowed region trivially, as it would mean that the sum of number of incomparable pairs and the number of links is greater than the total number of pairs.

3.3.3 Two-Layer Posets

For two-layer posets, we can choose $f_1 = f$ and $f_2 = 1 - f$. The equations simplify to:

$$\rho_0 = f^2 + (1-f)^2 + 2(1-\alpha)f(1-f) \tag{3.55}$$

$$\rho_1 = 2\alpha f(1-f) \tag{3.56}$$

We can see that $\rho_0 + \rho_1 = 1$. This is obivious because if the elements are comparable, they must be related by just links in a two-layer poset. The minimum of ρ_0 is $1/2$ at $\alpha = 1$ and $f = 1/2$,

and at the same point we obtain the maximum of ρ_1 as $1/2$.

We obtain the maximum value of entropy when $\alpha = 2(1-\rho_0)$ and $f = 1/2$, just like in the previous section. We now observe that we can think of α as $2\rho_1$ too. This implies that the maximum entropy is obtained when $\rho_0 = 3/4$ and $\rho_1 = 1/4$.

3.3.4 Three-Layer Posets

Choosing $f_1 = x$, $f_2 = y$ and $f_3 = 1-x-y$, we have the equations:

$$\rho_0 = x^2 + y^2 + (1-x-y)^2 + 2(1-\alpha)(xy + y(1-x-y)) \tag{3.57}$$

$$\rho_1 = 2\alpha(xy + y(1-x-y)) \tag{3.58}$$

Seeing that $\rho_1 = 2\alpha y(1-y)$, we conclude that $\max \rho_1 = 1/2$ at $\alpha = 1$ and $y = 1/2$. We also note that the minimum of ρ_0 is $1/3$ when all three layers have equal fraction of elements and $\alpha = 1$.

Now, we think of the combination $\rho_0 + \xi\rho_1$ and try to find linear bounds on the region where solutions are possible.

$$\rho_0 + \xi\rho_1 = 2x^2 - 2x + 2xy + 1 + 2\alpha(\xi - 1)(y - y^2) \tag{3.59}$$

One can easily analyse the expression $f(x,y) = 2x^2 - 2x + 2xy + 1$ and find that it is a positive function with range $\left(\frac{1}{2}, 1\right)$, attaining minimum as $x \to 1/2$ and $y \to 0$, and maximum as $x + y \to 1$. The remaining part has a range in $\left(0, \frac{\alpha(\xi-1)}{2}\right)$.

Fixing $\xi = 1$, leads to the bounds $1/2 < \rho_0 + \rho_1 < 1$ and fixing $\xi = 1/2$ leads to the bound $\rho_0 + \frac{1}{2}\rho_1 > 1/2$.

There exist some interestings points in this region. At $\left(\frac{1}{3}, \frac{4}{9}\right)$, the entropy is zero since only configuration (with all layers with probability $1/3$ and $\alpha = 1$) is allowed. Any line where $\alpha = 1$, the entropy is zero.

Consider a three-layer poset in which $y = 1/2$. It can be shown that in this case, for $\alpha = 1/2$, the entropy is maximum at $\frac{1}{2}\ln 2$ from $\rho_0 = 5/8$ to $\rho_0 = 3/4$.

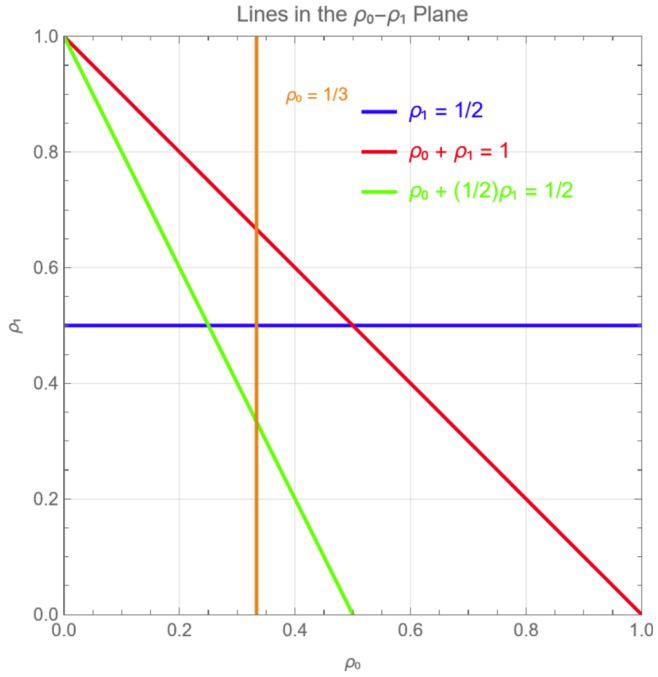


Figure 3.3: Linear bounds on the region in the $\rho_0 - \rho_1$ plane where the entropy function is well-defined for three-layer posets. A check for existence of solutions in this region shows similar but stronger bounds.

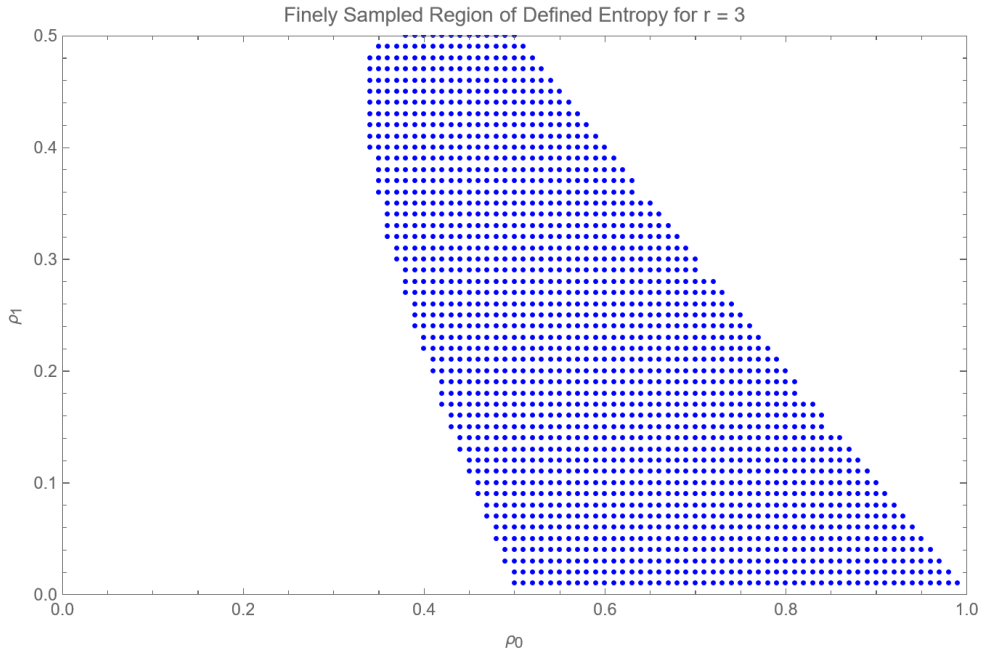


Figure 3.4: Region of the $\rho_0 = \rho_1$ plane where some defined entropy exists.

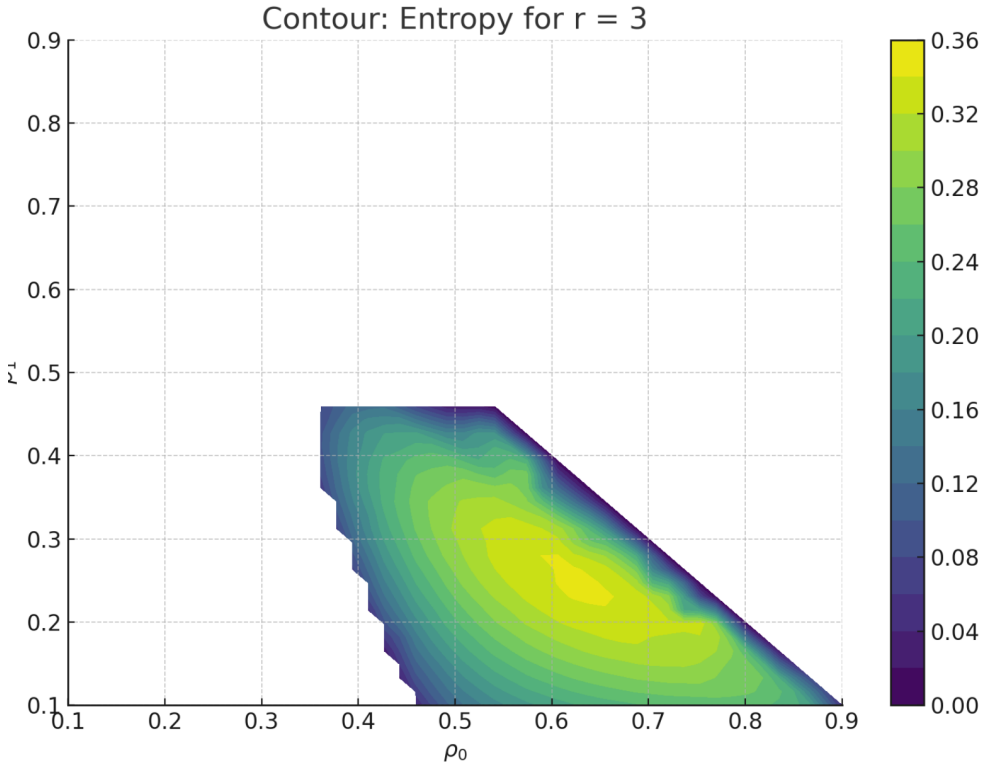


Figure 3.5: Contour plot of the entropy function $S(\rho_0, \rho_1)$ for three-layer posets. Cutoffs for ρ_0 and ρ_1 are put to avoid edgse cases. Observe that, apart from numerical tolerance, the maximum entropy behaves like $\frac{1}{2}\ln 2$ just like expected.

3.3.5 Higher Layer Posets

An analytic bound on the region in the $\rho_0 - \rho_1$ plane for higher layer posets is difficult to obtain. Numerical data suggests that the region increases. This means that one can obtain a layered poset with a large number of layers for the same combination of ρ_0 and ρ_1. The entropy function is continuous in this region, and the phase transitions are first order again. The curves where regions of r and $r+1$ layers meet are not the critical points though!

This can be illustrated by the case when $y = 1/2$ in a three-layer poset. Then, the minimum value of $\alpha = 2\rho_1$. Since ρ_1 is less than or equal to $1/2$, we obtain that

$$\rho_0 + \rho_1 = x^2 + \left(\frac{1}{2} - x\right)^2 + \frac{3}{4} \in \left(\frac{7}{8}, 1\right) \tag{3.60}$$

In this region, where $\rho_0 \in (3/8, 1/2)$, we find that the entropy follows exactly the same function as the two-layer poset, i.e. $\frac{1}{2}\mathcal{S}(2\rho_1)$. Hence, the critical lines are difficult to determine in general. We can say that the transition between two and three layers occurs somewhere in the region $\rho_0 \in (3/4, 7/8)$, and some corrsponding range of ρ_1.

Part II

Percolating Models

Chapter 4

Percolation Theory

It has been a long journey from the first chapter to this one. We have the complete picture of a spacetime if we know what event happens before or after another, and that this is equivalent to a partial order. A Lorentz covariant kinematics on a poset obtained by random sampling has been established. In order to completely characterise the kinematics of a causal set, we need to assign statistical weights to each element in a sample space of posets. The first step towards this has been taken in the previous chapter, where we have tried to count the number of posets. Having done that, we obtained that there are first order phase transitions in the entropy of a poset as a function of the density of chain lengths.

All this may seem very scattered. To bring this together coherently, we now discuss something called *percolation*. This may seem like a haphazard jump.

4.1 Introduction

Chapter 5

Models for Spacetime

I have tried to convince the reader that specifying the kinematics of a causal set is a humongously difficult task, particularly because we need to assign weights to each element in a sample space of posets in order to obtain manifold-like causal sets. We have seen what percolation is, and we have seen that a directed percolation model defines a partial order on the lattice.

Our goal for this final chapter is to develop candidates of spacetimes that result from percolation. This gives two benefits; 1. layered posets arise naturally from percolation, leading to a possible relation between this and random sampling, 2. the randomness in the kinematics is contained in a (bond/site depending on the model) probability instead of sampling from a well-defined manifold.

5.1 Directed Percolation on a Square Lattice

Consider a $\mathbb{Z}^2$ lattice with each bond having a probability $p > p_c$ of being occupied. Paths from a point x to a point y are defined only if there is a directed path (up and right only) from x to y in the lattice. Let's say we start the percolation at a minimal element which we mark as our origin. Averaged over many realizations, we expect there to exist an opening angle θ as a function of $p - p_c$ which governs the future and past of the origin. We observe that this looks like a light cone, albeit a generalised one. We try to map to this light cone to a Minkowski light cone by defining a Lorentzian metric on the lattice. If we can do this successfully, we can say that we have found a model for 1+1 dimensional special relativity. To do this, we need to find a reasonable definition of proper time on the lattice in this model. Of course, we expect these ideas to hold in the far future and far past only.

Appendix A

Causal Spaces

A.1 Filling the gaps in Kronheimer and Penrose

A.2 Examples of Causal Spaces