Topics in performance analysis

Eitan Bachmat
Contents

1. Introduction 5

Chapter 1. A classical model for storage system activity 9
1. Modeling disk drives 9
2. Modeling access patterns 11
3. Estimating seek times in the IRM model 14
4. Metric spaces 16
5. The IRM seek estimate is pessimistic 18
Notes for chapter 1 24

Chapter 2. A fractal model for storage system activity 31
1. Definition of the model 32
2. Average seek distance calculations 34
3. Cache hit ratios and entropy 36
4. Asymptotic hit ratio computations 39
5. Incompatibility of $q$-ary and $p$-ary bias models 42
Notes for chapter 2 47

Chapter 3. Disk scheduling, airplane boarding and Lorentzian geometry 51
1. The disk scheduling problem 51
2. Partially ordered sets 52
3. The ABZ algorithm 54
4. The general airplane boarding process 59
5. Disk scheduling and airplane boarding in a probabilistic setting 61
6. Lorentzian geometry 64
7. Discrete space-time 78
8. Modeling disk scheduling and airplane boarding via Lorentzian geometry 91
9. Computing the maximal curve 93
10. Analysis of back-to-front boarding policies 97
11. Disk scheduling under the microscope 104
12. Lorentzian geometry and general relativity theory 108
Notes for chapter 3 109

Chapter 4. Mirrored configurations 113
1. Modeling a mirrored data configuration using graphs 113
2. An upper bound on the cycle threshold 116
3. Lubotzky-Phillips-Sarnak (LPS) graphs 120
4. Constructions for general degrees and graphs sizes 132
   Notes for chapter 4 137

Chapter 5. On queues and numbers 141
1. Managing a mini-market 142
2. Basic queueing theory 150
3. Variants of the M/G/1 queue 156
4. Duality theory 161
5. Distributions in queueing theory 164
6. Modular forms 167
7. Hecke newforms and the mini-market theorem 179
8. SITA queues with Pareto job size distributions 183
9. Analysis of the large h asymptotics of SITA queues 190
   Notes for chapter 5 193

Chapter 6. Appendix A: Some basic definitions and facts 201
1. Point-set topology and metric topology 201
2. Measure theory 206
3. Probability 209
4. Algebra 212
5. Linear algebra 213
6. Harmonic analysis and transforms 214
7. Complex analysis 217
8. Basic definitions in graph theory 217

Chapter 7. Appendix B: Proofs of theorems 221
1. Lorentzian geometry 223

Bibliography 239

Index 249
1. Introduction

This book is about various topics in performance analysis. The selection of topics is extremely subjective and involves exclusively topics which I happened to work on in the last 15 years. The other selection criterion was that the topic actually had to involve some reasonably interesting mathematics. Thus some work which was technologically very relevant, but mathematically trivial, did not make the cut.

I did try to be honest in the problem selection, rather than trying to artificially invent a performance analysis problem that will fit some math that I knew, I first learned about the problems from a performance analysis perspective, either through real product development or by reading some papers on the subject. In fact, I learned a lot of new math in the process of trying to solve the problems or better understand some papers.

The goal of producing an elementary text of reasonable length that is suitable for advanced undergraduate/beginning graduate students means that full proofs are not always possible. However, reading the literature in the bibliography will fill those gaps. In fact, a major motivation for the book is to encourage the readers to explore that literature. With that goal in mind, we end each chapter with notes that point in related directions, putting the specific problem we considered in a larger perspective. We also provide at the end of each chapter some related research questions at various levels. Most are probably worth a research paper or can form the basis for a thesis.

Chapter 1 considers a problem which came up during the design of a commercially available software system that dynamically changes the configuration of a storage system in order to improve performance, [10]. The software could only use very coarse data and the problem was to guarantee that it will not make bad decisions as a result.

Chapter 2 is the result of trying to better understand [183], a paper which I really like. The authors introduced a very nice model for user access patterns and showed that it captures many of the characteristics of actual user access patterns. They explored the model empirically and made some insightful statements. We show how the model can be studied analytically and how to actually prove the statements.

Chapter 3 involves an attempt to analyze a practically important performance feature of disk drives. They can sometimes re-order I/O requests to improve service time. While our attempt to analyze the performance advantages of this feature ultimately failed, we can solve a toy model. Following a suggestion of Steven Skiena, the toy model can be used to model airplane boarding as well. The model seems to do a somewhat better job at modeling airplane boarding than the original scheduling problem. In a strange twist, it turns out that the same math involved in the modeling, has already been used to model the universe.
Chapter 4 deals with the problem of optimally configuring a storage system where each piece of data has two copies for fault tolerance purposes. This scenario can be used to improve the performance of data read commands since we can choose which copy will service the request. The problem of choosing a good data layout was raised by Yoav Raz at a briefing in which I explained the design of a software tool which was developed to manage read requests in such a system.

Chapter 5, the only one which is not about storage systems, is the product of an attempt to better understand a couple of very nice papers in queueing theory, [85],[73]. Queueing theory is a staple of performance analysis, which deals with queues such as the ones we may experience in supermarkets or banks. Specifically, we try to analyze a mini-market with two checkout counters, one serving as an express checkout line.

As the alert reader may have noticed, most of the chapters involve the performance of storage systems. This is not the first book to mathematically analyze storage systems. Two inspiring books on the subject have already appeared. The first, [188], by Wong, uses mostly combinatorial methods. Apart from disk based systems it analyzes automated tape libraries and drum based storage. The second, [13], by Aven, Coffman and Kogan, uses stochastic methods and differential equations to study cache, drum storage and disks. The problems we consider are entirely different, from those studied in these monographs, partly because the technology issues have changed.

On the subject of queueing theory there are many excellent books. My favorite basic introduction is [113]. Apart from storage and queues, performance analysis deals with many other systems including communications systems and networks. The reader may want to consult the proceedings of the annual conference SIGMETRICS which is devoted to the subject.

I have briefly described the problems but I did not give hints about the mathematical techniques involved. Well, the story is as follows.

As a teen-ager, I started reading a hard cover copy of Feynman’s lectures on physics, [74] which my dad bought about a decade before. After reading the last of the three chapters on special relativity theory, titled ”Space-Time” (chapter 17 of volume I), I skipped immediately to the last chapter (42) of volume II which is titled ”curved space” and explains the essentials of general relativity. I considered the content of these chapters to be the most exciting thing I have ever read and I immediately decided that I wanted to be a physicist.

At the university, it was obvious in the first year of studies that I was not talented enough to do physics and I abandoned the subject. I did somewhat better in math and I really liked it, so I decided to make an adjustment and pursue math as my main subject.

I received an excellent mathematical education in all my degrees. As an undergraduate at Hebrew U., I learned about the Gamma function and Stirling’s formula from Joram Lindenstrauss, about the zeta function and Dirichlet’s work from Hillel Furstenberg. Hillel also taught me about the
various transforms in probability and harmonic analysis. Benji Weiss taught me ergodic theory including the ergodic theorem, the sub-additive ergodic theorem and about entropy. Benji also taught me point-set topology with all its different spaces, metric or otherwise and with the corresponding notions of convergence. I learned many other basic tools of the trade from many other excellent teachers which are too numerous to mention.

In my Master’s degree, again at Hebrew U., I worked under the supervision of Joram Lindenstrauss on problems of embedding metric spaces into Hilbert space. Alex Lubotzky taught me about the newly found Lubotzky-Phillips-Sarnak graphs. I also participated in a number theory course by Peter Sarnak who was visiting Hebrew U. at the time. Among other things modular forms and the zeta function were discussed. During this period I also became interested in algebraic geometry, reading Hartshorne’s classical book on the subject, [91].

I did my PhD at M.I.T. under the supervision of Alexander Beilinson where I learned about arithmetic geometry, number theory, L-functions and their special values.

After completing my PhD I became a post-doc at Brown. I am sure I was the worst post-doc in the history of the Brown math department. I realized that I am not good enough to be a professional mathematician. I loved math but it did not love me back. It was time to make an adjustment again.

Around that time, I started consulting in performance analysis to a company that produces disk arrays, which are large computer storage systems. I knew absolutely nothing about performance analysis nor about storage systems, however, people like Natan Vishlitzky and Moshe Yanai, giants of storage systems technology, taught me the ropes. I started working with a fellow mathematician Dan Arnon and we were doing well, giving sound advice and helping the engineers improve system performance, I finally found something I was good at!

However, none of my work seemed to use any math whatsoever. It was based almost exclusively on some practical, common sense, attitude. My main advantage over the engineers was simply that I only had to give advice, while they actually had to implement the ideas, so I had more time to think about them.

I was sad that I was losing touch with mathematics, probably forever, and that all the beautiful stuff that I learned will slowly gather dust in my memory and fade into oblivion.

However, a miracle occurred. Slowly but surely, performance analysis problems that required the mathematics that I mentioned above started popping up in different circumstances. In fact, I had to learn some new and exciting math in order to deal with them. Ofer Zeitouni became my probability theory guru and Percy Deift opened up other exciting vistas. I learned a lot about performance analysis itself, from discussions with Ed Cofmann, Leo Flatto, David Gamarnik and especially Mor Harchol-Balter.
I also learned from reading some of their papers. The math made my work more academically viable, renewing my academic career, this time as a performance analysis researcher. It also made the work much more gratifying.

This book is a tribute to all the people who taught me math and is intended to share my excitement about how all this math can turn out to be useful in surprising ways.

Chapters 1, 3, 4, and to some extent, chapter 5 as well, are based on various papers that I co-authored, [19, 15, 16, 17, 5, 20]. I would like to thank all my colleagues who participated in these research projects, and in particular Noga Alon, Dani Berend, Tao-Kai Lam, Luba Sapir, Steven Skiena and Hagit Sarfati.

I would like to pay special tribute to Avner Magen who made crucial contributions to the material in chapter 1. Avner tragically died in May 2010 in a mountaineering accident. I hope this book contributes in a modest way to preserving the legacy of this wonderful person.

Finally I want to thank my parents Yehuda and Menuha and last but certainly not least, my wonderful family, Hagit my wife, and my kids Katya, Zoe and Daniel without whom life would simply be a long march.
A classical model for storage system activity

In the first four chapters of this book we will attempt to analyze the performance of storage systems. In reality, storage systems and the way they are used by applications are messy objects which are hard to describe. We therefore need to come up with idealized models, which are mathematically more negotiable, yet reflect some realistic system aspects. This chapter will introduce a basic model of a disk drive. We also introduce a very basic model for how the disk drive is used by applications. We then study some basic properties of these models.

1. Modeling disk drives

Disk drives are currently the main physical devices for storing computer data. We begin by mathematically modeling the physical characteristics of a disk drive. This model will serve us in the first three chapters of the book. The disk drive contains data on a platter, which is mathematically an annulus. The platter rotates at a constant speed. In addition to the platter that contains the data, a disk drive has a point-like read/write head which hovers over the platter. The read/write head of the disk can only move in, towards the center, and out, towards the edge, radially. When the head has to read or write data, it moves radially to the radial position of the data and then, waits for the rotation of the platter to bring the data underneath it. Mathematically, it is more convenient to describe this system by thinking of the platter as fixed and of the read/write head as rotating at a constant rate. In the following subsection we discuss the various ingredients that we use for mathematically modeling the behavior of a disk drive.

1.1. Coordinates for data locations. We model the disk drive platter in polar-like coordinates. For convenience, the radial coordinate will have range $0 \leq r \leq 1$ and similarly, we let the angular coordinate have range $0 \leq \theta \leq 1$, instead of the more standard range in $[0, 2\pi]$. The annulus is formed via the identification of points $(r, 0)$ with $(r, 1)$. Note however, that unlike polar coordinates we do not identify the points $(0, \theta)$, since the platter is an annulus. We can conveniently normalize our time parameter $t$ so that at time $t = 0$, the angular position of the head of the disk is $\theta = 0$, or in other words, $\theta(0) = 0$. We can also normalize the time parameter so that the disk platter makes a full rotation after precisely one time unit. Given that the platter rotates at a constant speed and that the head only moves
radially, this implies that
\[ \theta(t) = t \mod 1 \]

1.2. Seek time functions.

Definition 1. The seek time function \( f(r) \) measures the time required for the head to radially move a distance \( r \). The motion must start and end without radial velocity since the disk head must be stable to read and write data at a given radius.

The seek distance function \( g(t) \) is defined as the inverse (as a function) of the seek time function. It tells us how much distance the head covers, moving radially for \( t \) time units.

Due to effects of acceleration and deceleration, the seek function is always concave, namely
\[
f(cr_1 + (1-c)r_2) \geq cf(r_1) + (1-c)f(r_2)
\]
for any \( 0 \leq c \leq 1 \) and \( 1 \geq r_1, r_2 \geq 0 \). Geometrically this means that the graph of \( f \) lies above, or on, the line segment connecting any two points of the graph. For a function \( f \) which is twice differentiable concavity is equivalent to the statement that \( f''(r) \leq 0 \) for all \( r \geq 0 \). The seek time function allows us to make the following definition. The seek time function is also strictly increasing. Since the seek time function is concave it is easy to check that the seek distance function is convex, i.e.,
\[
g(ct_1 + (1-c)t_2) \geq cg(t_1) + (1-c)f(t_2)
\]
for any \( 0 \leq c \leq 1 \) and \( f(1) \geq t_1, t_2 \geq 0 \).

Definition 2. Let \( L_1 = (r_1, \theta_1) \) and \( L_2 = (r_2, \theta_2) \) be two locations on the disk drive. The seek time between \( L_1 \) and \( L_2 \) is defined as
\[ f(|r_2 - r_1|) \]
where \( f \) is the seek time function of the disk drive. This is the time required by the disk head to move radially from radius \( r_1 \) to radius \( r_2 \).

If the head of the disk is reading or writing data at time \( t_1 \) at location \((r(t_1), \theta(t_1))\) and at time \( t_2 \geq t_1 \) is reading or writing data at location \((r(t_2), \theta(t_2))\), then \( t_1, t_2 \) must satisfy the seek time constraint relation
\[ t_2 - t_1 \geq f(|r_1 - r_2|) \]

After seeking to radial position \( r_2 \), the head waits without radial movement for the data at angle \( \theta_2 \) to pass underneath it. The constant speed rotation of the disk which is responsible for equation (1), together with the seek time function constraint (2) on read/write times and locations, provide the following mathematical characterization of the disk drive and the motion of its head.


2. MODELING ACCESS PATTERNS

Definition 3. Let $L_1 = (r_1, \theta_1)$ and $L_2 = (r_2, \theta_2)$ be two locations on a disk drive with seek time function $f$. We define $d_f(L_1, L_2)$, the time required to move from location $L_1$ to location $L_2$ as the minimum of $t_2 - t_1$ taken over all pairs $t_1, t_2$ such that $\theta(t_1) = \theta_1$, $\theta(t_2) = \theta_2$ and $t_2 - t_1 \geq f(|r_1 - r_2|)$.

In many situations we are only interested in radial aspects of the motion. In such cases we will model the disk not as an annulus, but more simply as a unit size segment $0 \leq r \leq 1$. The segment will be equipped with a concave seek function $f(|r_1 - r_2|)$ which measures the seek time between any pair of points $r_1, r_2$ in the segment.

We can also consider a discretized version. In this version we have a finite set of points $0 \leq r_1 < r_2 < \ldots < r_n \leq 1$. We assume that all requests are to locations whose radial coordinate is one of the $r_i$. The seek time becomes a function

$$d_{i,j} = f(|r_i - r_j|) \tag{3}$$

In this book we will use both the discrete and continuous models as needed. The present chapter will use a discrete model for the radial direction only.

2. Modeling access patterns

Having modeled the physical (hardware) aspects of a disk drive, we need to model the way in which it is used by software applications. Computer programs use the disk by sending to it requests to read or write data in various locations. Such requests are known as I/O (input/output) requests. For our purposes it will not matter much whether the request is a read or write request. A request will therefore be specified by a time $t$ in which the request was made and a disk location $s = (r, \theta)$ which tells the disk where to read or write.

Definition 4. A sequence of requests represented by pairs of request times and disk locations, $(t_i, s_i)$, $i = 1, \ldots, a$, satisfying $t_i < t_j$ for $i < j$ (ordered by time), is called a trace. We say that $a$ is the length or size of the trace.

In some cases it is convenient to model just one of the aspects of the trace, either time or space, or even more restricted data such as radial locations of the requests. For example, if we wish to study seek times, and the requests are serviced in the order of their arrival, then it is enough to know the radial locations of the sequence of requests. If we assume a discrete model of disk locations $r_1, \ldots, r_n$ the sequence (trace) of radial disk locations will have the form $r_{i_1}, \ldots, r_{i_a}$. The total amount of time that the disk spent seeking between requests in the trace is $\sum_{i=1}^{a-1} d_{i_1, i_{i+1}}$, where the function $d$ is given by (3).

Mathematically, it is more convenient to add one more (virtual) seek between the final location $r_{i_a}$ and the first $r_{i_1}$, closing a cycle. The total
seek time between requests in the cycle, which we denote by $T$ is given by

$$T = \sum_{i=1}^{a} d_{i,i+1}$$

where by definition $r_{i+1} = r_i$. Usually, in applications of such formula $a$, the total number of requests, is rather large so the addition of a single "virtual" seek has little practical effect. However, it will make the math much more elegant. Knowing the total seek time of a disk on a given trace is useful since it gives us an indication of how hard the disk is working.

A major problem with traces is that they are very rarely available. Commonly we are only given some statistics about the trace. A typical example is a set of counters which count the number of times that each radial location was requested with no information on the order or timing of the requests. Therefore, in order to analyze total seek time, or other quantities of interest, we would like to construct synthetic traces which are based on mathematical models which incorporate the statistical features of the real trace.

A second issue with real sequences of requests is that they are not very structured. Therefore it is hard to deduce from the particular behavior of one trace, how another will perform. Mathematical models with multiple parameters can model diverse situations by varying their parameters, they are therefore, much more flexible than any given trace.

2.1. The IRM model. Our first example of a mathematical model for generating traces is called the independent reference model (IRM). It will only model a sequence of radial disk locations, without reference to the exact timing of requests or their angular location. We will use it to estimate the total seek time. We will assume the discrete model with radial locations $r_1, \ldots, r_n$.

**Definition 5.** Let $X_P$ be a random variable such that $Pr(X_P = i) = p_i$. Let $P = (p_1, \ldots, p_n)$ be the associated probability distribution. An Independent Reference Model (IRM model), with probability distribution $P = (p_1, \ldots, p_n)$ is a method for generating traces of radial disk locations. To generate a trace of length $a$, we i.i.d. sample $X_P$ $a$ times to obtain the trace $r_{i_1}, \ldots, r_{i_a}$. In other words, for a fixed value of $a$, the IRM model assigns the probability

$$\prod_{j=1}^{a} p_{i_j}$$

to the trace sequence $r_{i_1}, \ldots, r_{i_a}$.

In order to use the IRM to model an actual trace we need a method for choosing the probabilities $p_i$ based on an observed trace sequence. Since we are assuming that the requests were made independently, the only thing that matters statistically, is how many times each location was requested.

Assume that we observed that the radial location $r_i$ was requested $a_i$ times out of a total of $a = \sum_i a_i$ requests. It would be intuitively natural
to let $p_i = a_i / a$, be our probabilities for the IRM. We denote this probability distribution by $P$. We can provide some more formal justification for this choice by considering the probability of observing the trace given an IRM model with a probability distribution $Q$ given by $q_1, \ldots, q_n$. The probability of observing a trace with $a_i$ occurrences of $r_i$ in an IRM model with distribution $Q$ is $Pr(Q; Trace) = \prod_i q_i^{a_i}$. The following well known proposition states that among all possible IRM models, the one with probability distribution $p_i = a_i / a$ maximizes the probability of observing the actual trace.

**Proposition 1.** For any probability distribution $Q$, we have

$$Pr(P; Trace) \geq Pr(Q; Trace)$$

and equality holds if and only if (iff) $P = Q$.

**proof:** It is easier to consider $\log(Pr(Q; Trace)) = a \sum_i p_i \log(q_i)$. Taking the difference and dividing by $a$ we get

$$\Delta = \frac{1}{a} (\log(Pr(Q; Trace)) - \log(Pr(P; Trace))) = \sum_i p_i (\log(q_i / p_i))$$

It is easy to check the general inequality $\log(x) \leq x - 1$. In fact the two sides are equal at $x = 1$ and the derivative of the difference is negative for $x < 1$ and positive for $x > 1$. This argument also shows that equality occurs only at $x = 1$. We apply this inequality with $x_i = q_i / p_i$ to obtain

$$\Delta = \sum_i p_i \log(x_i) \leq \sum_i p_i (\frac{q_i}{p_i} - 1)$$

$$= \sum_i (q_i - p_i) = \sum_i q_i - \sum_i p_i = 1 - 1 = 0$$

We also note that $\Delta = 0$ iff for all $i$, $q_i / p_i = 1$. Together these observations prove the proposition. **q.e.d.**

Before proceeding, it is interesting to note that during the proof of the proposition we have encountered some famous and important quantities.

**Definition 6.** Given a discrete probability distribution $P = (p_1, \ldots, p_n)$, the quantity

$$H(P) = - \sum_i p_i \log(p_i)$$

is called the **entropy** of the distribution $P$. Given two distributions $P, Q$ the quantity

$$D(P||Q) = - \sum_i p_i (\log(q_i / p_i)) = \sum_i p_i (\log(p_i / q_i))$$

is called the **relative entropy** of $P$ w.r.t. $Q$. It is also known as the Kullback-Leibler distance.
The proof of the proposition shows that $D(P||Q) \geq 0$ and that equality holds iff $P = Q$. The entropy measures in bits how much information on average is gained by sampling the probability distribution once. Both the entropy and the relative entropy play a major role in mathematics and physics. The concept of entropy originates in the study of thermodynamics. It is also especially important in the field of information theory, which provides a mathematical basis for the design of communication systems. We will encounter these concepts again in chapter 2 of the book.

3. Estimating seek times in the IRM model

We have seen how to set parameters for the IRM from an observed sequence of activity $a_i$. We would like to compute an estimate of a meaningful measure of disk performance. We consider the average seek time between consecutive radial locations in a trace.

**Proposition 2.** The average seek time between consecutive requests which are generated by the IRM with probability distribution $P$ is

$$
\sum_i \sum_j p_i p_j d_{i,j}
$$

**Proof:** The probability that a request will be to location $r_i$ and that the next request will be to location $r_j$ is by independence $p_i p_j$. If that happens the seek time is $d_{i,j}$ and summing up over all possibilities gives the required calculation. $q.e.d.$

The proposition provides us with the average seek time of a single request. The average total seek time for all requests in a trace of length $a$ generated by the IRM model, $E(T)$, will be given by

$$
E(T)(d_{i,j},a_i) = a(\sum_i \sum_j p_i p_j d_{i,j}) = a(\sum_i \sum_j \frac{a_i a_j}{a} d_{i,j}) = \frac{1}{a} \sum_i \sum_j a_i a_j d_{i,j}
$$

This quantity will serve as our total seek time estimate.

We seem to be making good progress. We started with a trace, with some total seek time $T$, replaced it by an appropriate model and obtained a short formula for the total seek time estimate $E(T)(d_{i,j},a_i)$ using the model. One major problem is that not all traces are properly modeled by the IRM, so our estimate may be way off. In addition, our estimate is just a statistical average, so even for an IRM generated sequence it will not yield the correct result. Let us consider an example.

**Example:** Consider a disk with just two radial locations $r_1$ and $r_2$ and let’s assume that the seek time between the locations is $d_{1,2} = 1$. We consider two different traces. Trace A consists of 1000 consecutive requests to $r_1$ followed by 1000 requests to $r_2$. Trace B consists of 1000 alternating
calls to \( r_1 \) and then \( r_2 \), leading to the index sequence 1, 2, 1, 2, 1, 2, \ldots. In both cases there were 1000 calls to \( r_1 \) and 1000 calls to \( r_2 \). We conclude that both traces will be modeled by an IRM with probabilities \( p_1 = p_2 = 1/2 \). Plugging into equation (6) we obtain a total seek time estimate of 1000. On the other hand trace A had two seeks between radial locations, one in the middle of the trace and one in the end to close the cycle, leading to a total seek time of 2, while trace B had a total seek time of 2000.

We see that our estimate was off in both cases, badly over-estimating the total seek of trace A while under-estimating the total seek of trace B by a factor of 2. Using essentially the same example but with more requests shows that the IRM can over-estimate the total seek time by an arbitrarily large factor and can under-estimate by at least a factor of 2.

Can we produce examples where the under-estimation factor is even worse?

This is an important question. Total seek time is a measure of how hard a disk is working. If we overestimate this measure we may be led to falsely conclude that a disk is working hard when in reality it is not. While this may be inconvenient, the other type of error, under-estimation is far worse for diagnostic purposes. Under-estimation can lead us to falsely believe that a disk is nearly idle, when in fact it is working hard. Since seeks are time consuming they usually result in inferior performance, if we under-estimate this problem we may ignore it, instead of taking corrective measures. It is the same issue as with diagnosis of a disease, where over and under estimation are not symmetrical.

We can rephrase the question more mathematically as follows. Let \( W(d_{i,j}, a_i) \) be the maximal value of the total seek time \( T \) over the set of all possible request sequences \( r_{i_1}, \ldots, r_{i_n} \), with \( r_i \) appearing \( a_i \) times in the sequence. If our counters show that \( r_i \) was requested \( a_i \) times then obviously the actual trace belongs to this set and therefore, the actual seek time \( T \) satisfies \( T \leq W(d_{i,j}, a_i) \).

The rephrased question is whether we can find examples where the IRM estimate \( E(T)(d_{i,j}, a_i) \) is less than \( 1/2 W(d_{i,j}, a_i) \). Our first result answers the question negatively, but first, we need to introduce an important concept.

**Definition 7.** A pair \((X, d)\) where \( X \) is a set and \( d \) is a function \( d : X \times X \rightarrow \mathbb{R}^+ \) is called a metric space if

(A) \( d(x, x) = 0 \) for all \( x \in X \) and \( d(x, y) > 0 \) for all \( x \neq y \).
(B) \( d(x, y) = d(y, x) \) for all \( x, y \in X \).
(C) \( d(x, y) + d(y, z) \geq d(x, z) \) for all \( x, y, z \in X \).

If instead of property (A) we only require that \( d(x, x) = 0 \) we say that \((X, d)\) form a semi-metric.
Upon examination we can easily verify that seek times given by equation (3), with \( f \) a concave function, do indeed produce a metric space structure on the set \( X \) of radial disk locations. In fact, we get a metric space structure in the continuous setting as well.

**Theorem 1. (Comparison of the IRM estimate and the actual seek time)**

If \( d = d_{i,j} \) is a metric on the radial disk locations then for any set of multiplicities \( a_i \) we have

\[
2E(T)(d, a_i, IRM) \geq W(d, a_i)
\]

**Proof:** Assume first that the multiplicities \( a_i \) are all equal to 1. In this case the IRM estimate is \( \frac{1}{a} \sum_{i,j} d_{i,j} \). Assume without loss of generality that the maximal value of \( T \) on sequences with multiplicities \( a_i = 1 \) is given by the sequence \( 1, 2, \ldots, a \). Since \( d \) satisfies the triangle inequality we have for \( 1 \leq i < a \) and for disk location index \( j \), the inequality, \( d_{i,i+1} \leq d_{i,j} + d_{j,i+1} \) (the \( a + 1 \) location index coincides with the first). Summing over all \( i,j \) we get

\[
a \sum_{i=1}^{a} d_{i,i+1} \leq 2 \sum_{i,j} d_{i,j}.
\]

Therefore \( 2E(T)(d, a_i) \geq W(d, a_i) \) in this case. To complete the proof we need to consider general multiplicities \( (a_1, \ldots, a_n) \). Let \( X' \) be the set of radial locations, with each location represented \( a_j \) times. Given \( d \) we construct a semi metric on the set \( X' \) by letting the distance between an element representing location \( r_i \) and an element representing location \( r_j \) be \( d_{i,j} \). Sequences of location indexes with multiplicities \( a_i \) are in 1-1 correspondence with sequences with multiplicities 1 in the new indexes. If we visit location \( i \) for the \( k' \)th time in the original sequence, we visit the \( k' \)th element representing location \( r_i \) in the corresponding sequence. By construction, this correspondence also preserves the seek time \( T \). Clearly \( d_{i,j} \) of the new construction also satisfies the triangle inequality. We have thus reduced the problem to the case where the multiplicities are all 1. q.e.d.

We have established one simple but very useful property of the IRM estimate. We can establish another nice property if we delve a little deeper into the theory of metric spaces.

### 4. Metric spaces

Perhaps the most natural example of a metric space is given by \( n \)-dimensional Euclidean space \( \mathbb{R}^n \), with the usual distance function \( d_2(x, y) = \sqrt{\sum_{i=1}^{n} \|x_i - y_i\|^2} \), where \( x = (x_1, \ldots, x_n) \) and \( y = (y_1, \ldots, y_n) \) are vectors.

**Definition 8.** We say that a metric space \((X, d)\) is **Euclidean** if there is an \( n \) and a map \( \phi : X \to \mathbb{R}^n \) such that for all \( x, x' \in X \) we have \( d(x, x') = d_2(\phi(x), \phi(x')) \), namely, \( \phi \) is distance preserving.
4. METRIC SPACES

A metric space \((X, d)\) is said to be **square Euclidean** if \((X, \sqrt{d})\) is Euclidean.

Not all metric spaces are Euclidean. As an example consider a metric space on 4 elements \(A, B, C, D\) given by \(d(A, B) = d(B, C) = d(C, D) = d(D, A) = 1\) and \(d(A, C) = d(B, D) = 2\). All other distances are determined by symmetry. We claim that \(A, B, C, D\) with the \(d\) metric is not Euclidean. Assume to the contrary that there is a map \(\phi\) of \(A, B, C, D\) into Euclidean space such that \(d_2(\phi(A), \phi(B)) = d(A, B) = 1\), \(d_2(\phi(B), \phi(C)) = d(B, C) = 1\) and \(d_2(\phi(A), \phi(C)) = 2\). These distance restrictions force \(\phi(A), \phi(B)\) and \(\phi(C)\) to be co-linear and moreover, \(\phi(B)\) is forced to be the mid point of the segment between \(\phi(A)\) and \(\phi(C)\). For the exact same reasons \(\phi(D)\) is also forced to be the mid point of the same segment, thus \(\phi(B)\) and \(\phi(D)\) must coincide, a contradiction.

The example above is square Euclidean. This can be seen by mapping \(A, B, C, D\) to the corners of the unit square in the plane, \(A, D\) being mapped onto opposite corners. In fact, it can be shown that any metric space with 4 elements is square Euclidean. However, it is possible to construct examples of metric spaces with 5 elements which are not square Euclidean.

There are two classical criteria for determining whether a finite metric space is Euclidean.

**Theorem 2. (Cayley’s criterion)** Consider a metric \(d_{i,j}\) on the set \(X = \{1, \ldots, n\}\). Consider the matrix \(M\) with entries

\[
M_{i,j} = \frac{1}{2}(d^2(i, n) + d^2(j, n) - d^2(i, j))
\]

with \(i, j = 1, \ldots, n-1\). Then, \((X, d)\) is Euclidean if and only if the matrix \(M\) is positive semi-definite, i.e., all of its eigenvalues are non-negative.

**Proof:** The matrix \(M\) is symmetric. By a basic result in linear algebra we can write \(M = ODO^t\), where \(O\) is an orthogonal matrix and \(D\) is a diagonal matrix whose entries are the eigenvalues of \(M\). If the matrix \(M\) is positive semi-definite then we can define the matrix \(\sqrt{D}\) to be the diagonal matrix with entries \(\sqrt{D}_{i,i} = \sqrt{D_{i,i}}\). The matrix \(B = O\sqrt{D}\) satisfies \(BB^t = M\). If we denote by \(x_i\), the \(i^{th}\) row of \(B\), then the scalar product of two rows satisfies \(x_i \cdot x_j = M_{i,j}\). We let \(\phi\) map the element \(n \in X\) to the origin of \(n-1\) dimensional Euclidean space and the element \(i < n\) to \(x_i\). Let \(d_2\) denote the Euclidean metric. For any \(i\) we have

\[
d_2^2(\phi(i), \phi(n)) = ||x_i||^2 = M_{i,i} = d^2(i, n)
\]

as required. For \(i, j < n\), we also have

\[
d_2^2(\phi(i), \phi(j)) = ||x_i - x_j||^2 = M_{i,i} + M_{j,j} - 2M_{i,j} = d^2(i, j)
\]

We can also run the argument backwards, starting with a mapping \(\phi\) choosing \(\phi(n)\) as the origin and obtaining a decomposition \(M = BB^t\), which shows that \(M\) is semi-positive definite. q.e.d.
Theorem 3. (Schoenberg’s criterion) A metric space \((X, d)\) is Euclidean if and only if for all real numbers, \(v_1, \ldots, v_n\), with \(\sum_v v_i = 0\) we have

\[
\sum_{i,j} v_i v_j d_{i,j}^2 \leq 0
\]

Proof: We assume that Schoenberg’s criterion holds. Let \(M\) be the matrix in Cayley’s criterion. It is a well known result in linear algebra that \(M\) is semi-positive definite if and only if we have \(wMw^t \geq 0\) for any \(n-1\) dimensional vector \(w = (w_1, \ldots, w_{n-1})\). Consider the matrix \(2M\) and convert it into an \(n\) by \(n\) matrix \(N\), by adding a last, zero row and column. we can write \(N\) as a sum of three matrices \(N_1, N_2, N_3\), given by \((N_1)_{i,j} = d^2(i,n), (N_2)_{i,j} = d^2(j,n)\) and \((N_3)_{i,j} = -d^2(i,j)\). This is obvious for the original \(n-1\) by \(n-1\) principal minor and a short verification shows that the last row and column of the sum are indeed zero. Given \(w\), we convert it into a zero sum, \(n\) dimensional vector \(v = (w_1, \ldots, w_{n-1}, -\sum_{i=1}^{n-1} w_i)\). We note that \(vN_2 = 0\), since the columns of \(N_2\) are constant entry vectors and \(v\) is a zero sum vector. Similarly, \(N_1 v^t = 0\) and therefore \(vN v^t = vN_3 v^t \geq 0\), the last inequality coming from the criterion. Running the argument backwards, yields the other direction. \(q.e.d.\)

5. The IRM seek estimate is pessimistic

We now go back to modeling. One of the basic features of worst case scenarios, or estimates, is that they only become more optimistic as more information is given. For example, someone might tell you that something bad happened to your friend. Given that information, the worst case scenario is that your friend is dead. You are then told that your friend is alive and at the hospital, that still leaves the possibility of a permanent coma, however, moments later you are told that the friend simply broke his/her foot. While that is still bad news, it is much better than what the previous worst case scenarios can lead you to believe. We have seen that the IRM model is close to being a worst case model. We can ask whether it has this property of being increasingly optimistic as more information is added.

Let \(I\) be a time interval and let \(I_1, \ldots, I_k\) be a sub-division of \(I\) into sub-intervals. For a time interval \(I\), let \(a_I\) be the vector whose \(i\)'th component is the number of times radial location \(r_i\) was requested during \(I\). Accordingly, we have \(a_I = \sum_{j=1}^{k} a_{I_j}\). Our estimate for total seek time during interval \(I\) will be \(E(T)(d, a_I)\). Once we are given the more detailed information \(a_{I_j}\), our new estimate for the total seek during time interval \(I_j\) is \(E(T)(d, a_{I_j})\) and the sum of such terms is the new estimate for time interval \(I\).
DEFINITION 9. We say that the IRM is super-additive with respect to a seek time function \( d \) if the inequality
\[
E(T)(a, d) \geq \sum_{j=1}^{k} E(T)(a_{I_j}, d)
\]
always holds.

Super additivity may be interpreted as stating that the addition of input information, namely, \( a_{I_j} \) instead of \( a_I \), never increases the estimate.

The rest of this chapter is devoted to proving the following result

**Theorem 4.** The IRM estimate for total seek time in a disk drive is always super-additive.

We have the following seek function dependent criterion for super additivity.

**Theorem 5.** (A criterion for super-additivity of the IRM) The IRM estimate is super-additive with respect to \( d \) if and only if \( d \) is square Euclidean.

**Proof:** It is enough to establish super-additivity for a subdivision of the time interval \( I \) into two sub-intervals, that is, to show that for all non-negative vectors \( a = (a_i) \), \( b = (b_i) \),
\[
E(T)(a + b, d) \geq E(T)(a, d) + E(T)(b, d)
\]
Let \( a = \sum_i a_i \) and \( b = \sum_i b_i \). Then
\[
E(T)(a + b, d) - E(T)(a, d) - E(T)(b, d) = -\frac{ab}{a+b} \sum_{i \neq j} d_{ij} \left( \frac{a_i}{a} - \frac{b_i}{b} \right) \left( \frac{a_j}{a} - \frac{b_j}{b} \right).
\]
Setting \( v_i = \frac{a_i}{a} - \frac{b_i}{b} \), we get
\[
(10) E(T)(a + b, d) - E(T)(a, d) - E(T)(b, d) = -\frac{ab}{a+b} \sum_{i \neq j} d_{ij} v_i v_j.
\]
We note that $\sum_i v_i = 0$, hence by Schoenberg’s criterion the IRM estimate is super additive if $d$ is square Euclidean. Conversely if the IRM estimate is super-additive then
\[ \sum_{i,j} d_{ij} v_i v_j \leq 0 \]
for all $v$ of the form $a/b - b/b$ where $a, b$ are vectors with integer non-negative entries. After scaling we may deduce that the property holds whenever $a, b$ have rational non-negative entries and by density of the rationals for all $a, b$ with non-negative entries. Every vector $v = (v_1, \ldots, v_h)$ such that $\sum_i v_i = 0$ has a multiple of the form $1/a - 1/b$, where $a, b$ have non-negative entries. Indeed if $a_i = \max\{v_i, 0\}$ and $b_i = \max\{-v_i, 0\}$, then $a = b$ and $\frac{1}{a} a - \frac{1}{b} b = \frac{1}{a} v$, hence Schoenberg’s criterion holds and $d$ is square Euclidean.

Following the theorem, the question of whether the IRM estimate is super-additive, becomes the question, whether the seek time metrics of disk drives are square Euclidean. We recall that the seek time of a disk drive has a special form $d_{i,j} = f(|r_i - r_j|)$ for a concave non-decreasing function $f$ with $f(0) = 0$. Let us denote a metric of this form by $d_f$.

In 1941, long before the arrival of disks, I.J. Schoenberg and J. Von Neumann, [180], gave a precise characterization of functions $f$ (concave or not) for which $d_f$ is square Euclidean for any choice of $r_i$. Using this characterization they showed that the functions $f(r) = r^\alpha$, $0 < \alpha < 1$ all have this property. However, it is hard to apply their theorem to general concave functions, thus it is not sufficient for our purposes. We describe their result in the notes at the end of the chapter.

The statement that for any non-decreasing concave function $f$ with $f(0) = 0$, the metric $d_f$ is square Euclidean was first proved by J. Kelly in 1972, [107], again without any relation to disk drives. In fact, Kelly proved a somewhat stronger statement. Combining Kelly’s theorem with theorem 5 shows that the IRM estimate of seek times in a disk drive is always super-additive.

To explain Kelly’s result and to prove an even stronger statement we introduce some more types of metric spaces. An important generalization of Euclidean space is obtained by considering the $\ell_p$ norms on $\mathbb{R}^n$, given by, $\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}$ where $p \geq 1$. For two vectors $x, y \in \mathbb{R}^n$, this defines a distance $d_p(x, y) = \|x - y\|_p$. The usual Euclidean distance is simply $d_2$.

**Definition 10.** A semi-metric space $(X, d)$ is called an $\ell_p$-metric if there exists a mapping $\phi : X \mapsto \mathbb{R}^n$ such that $d(x, y) = \|\phi(x) - \phi(y)\|_p$ for all $x, y \in X$.

We will be particularly interested in the case of $\ell_1$ metrics. To better understand $l_1$ metrics we consider the notion of a cut metric. These metrics are actually only semi-metrics.
Definition 11. Let $S \subset X$. The cut metric $d^S$ on $X$ associated with $S$ is the semi-metric given by $d^S_{i,j} = 1$ if either $i$ or $j$ are in $S$, while the other is not. Otherwise if both $i,j$ are in $S$, or both are not in $S$, $d^S_{i,j} = 0$.

Lemma 1. A semi-metric $d$ is an $l_1$ metric if and only if, it can be expressed as $d = \sum S c_S d^S$, where $c_S \geq 0$.

Proof: A cut metric is $l_1$, in fact consider the mapping $\phi : X \to \mathbb{R}$ which sends all elements of $S$ to 0, and all elements not in $S$ to 1. Obviously $d^S_{i,j} = d_{\phi(i),\phi(j)}$, where $d$ denotes the 1-dimensional Euclidean metric, which coincides with the 1-dimensional $l_1$ metric. We note that all $l_1$ metrics form a cone. Indeed, let $d^1,\ldots,d^s$ be $l_1$ metrics on $X$ and $\phi_i$ are the maps which show that the metrics are $l_1$. Let $d = \sum_{j=1}^s c_j d^j$, be some non-negative combination of the $d^j$ then the map $(c_1 \phi_1,\ldots,c_s \phi_s)$ which scales and then concatenates the maps $\phi_j$ shows that $d$ is also an $l_1$ metric. We conclude that all non negative combinations of cut metrics are $l_1$. To show the converse we first consider any set of points $x_1 < x_2 < \ldots < x_n \in \mathbb{R}$, with the usual metric $d_{i,j} = |x_i - x_j|$. We claim that the metric $d$ is a non-negative combination of cut metrics. Consider the subset $S_i \subset X$ consisting of $x_1,\ldots,x_i$, then it is easy to verify that $d = \sum_i (x_{i+1} - x_i) d^S_i$.

If $v^1,\ldots,v^n$ are $m$ dimensional vectors with coordinates $v^j_i$, $j = 1,\ldots,m$, with the $l_1$ metric $d$. we can consider the 1-dimensional metrics $d^1,\ldots,d^m$, where $d^j$ is the 1-dimensional metric on $v^j$, $i = 1,\ldots,n$. By the definition of $\ell_1$ we have $d = \sum_{j=1}^m d_j$ which shows that $d$ is a non-negative combination of cut metrics. q.e.d.

A more abstract class of metric spaces is provided by the hyper-metric inequalities.

Definition 12. We say that a metric space on a set of size $n$ is a hyper-metric if for all sequences of integers $x_1,\ldots,x_n$ such that $\sum_i x_i = 1$ we have

$$\sum_i \sum_j d_{i,j} x_i x_j \leq 0$$

We make a few remarks on this definition:

(1) One of the inequalities that a hyper-metric must satisfy is the triangle inequality. Specifically, the inequality $d_{i,j} + d_{j,k} \geq d_{i,k}$ is reproduced by setting $x_i = x_k = 1$, $x_j = -1$, and all other values equal to 0. This justifies the name hyper-metric as a particular type of metric.

(2) It is obvious from the definition that non-negative combinations of hyper-metrics are hyper-metrics.

(3) The condition that $x_i$ be integer is important. If we drop the requirement, then, $\sum_i \sum_j d_{i,j} x_i x_j \leq 0$ whenever $\sum_i x_i \neq 0$, which by continuity implies the inequality for all vectors $(x_1,\ldots,x_n)$. This
means that the matrix $d_{i,j}$ is negative semi-definite, which is equivalent to the statement that all its eigenvalues are non positive. However the trace of the matrix, which is the sum of all eigenvalues is $\sum_i d_{i,i} = 0$, which means that the matrix is 0.

(4) By applying the inequalities to the vector $-(x_1, \ldots, x_n)$ we see that we can replace the requirement that $\sum_i x_i = 1$ by the requirement that $\sum_i x_i = -1$. The following lemma links some of the classes of metric spaces which we introduced.

**Lemma 2.** We have the following sequence of containments between classes of metric spaces

$$\ell_1 \subset \text{Hypermetrics} \subset \ell_2^2$$

where $\ell_2^2$ denotes the space of square Euclidean metrics.

**Proof:** We first consider the statement $\ell_1 \subset \text{hypermetrics}$. By the previous lemma and remark 2 on hyper-metrics it is enough to show that any cut metric $d^S$ is a hyper-metric. Let $x_1, \ldots, x_n$ be a vector of integers with $\sum_i x_i = 1$. Let $x_S = \sum_{i \in S} x_i$, then,

$$\sum_i \sum_j d^S_{i,j} x_i x_j = 2 \sum_{i \in S} \sum_{j \notin S} 1 = 2x_S(1 - x_S) \leq 0$$

the last inequality since $x_S$ is an integer.

To check that a hyper-metric is in $\ell_2^2$ we must show according to Schoenberg's criterion that for any vector $(x_1, \ldots, x_n)$, such that $\sum_i x_i = 0$, we have $\sum_i \sum_j d_{i,j} x_i x_j \leq 0$. Assume first that all the $x_i$ are integers. We may produce from the $x_i$, a new vector $x'_i$ which satisfies the condition $\sum_i x'_i = 1$ by setting $x_i = x'_i$ for $i < n$ and $x'_n = x_n + 1$. By the hyper-metric assumption, applied to the vector $(x'_i)$, we get

$$\sum_i \sum_j d_{i,j} x'_i x'_j + 2 \sum_{i=1}^{n-1} d_{i,n} x_i \leq 0$$

By remark 4, we can also apply the hyper-metric inequality to the vector $x''_i$, given by $x''_i = x_i$ for $i < n$ and $x''_n = x_n - 1$. We obtain

$$\sum_i \sum_j d_{i,j} x_i x_j - 2 \sum_{i=1}^{n-1} d_{i,n} x_i \leq 0$$

Combining the inequalities leads to the desired conclusion for integer vectors. If the vector has rational coefficients we may scale it by its least common denominator and apply the inequalities to the resulting integer vector, so we get the desired result for all vectors with rational coefficients. By continuity we get the result for all real vectors as required. *q.e.d.*

Kelly proved that if $f$ is concave, non-decreasing with $f(0) = 0$ then $d_f$ is a hyper-metric for all values of $r_i$. 
Instead of proving Kelly's result, we will prove the stronger result that if \( f \) is concave, then, \( d_f \in \ell_1 \). Before we do so we need one more fact which we leave as an exercise. We can consider the space of functions \( g \) on \( \mathbb{R} \) such that \( \int_{\mathbb{R}} |g(x)| \, dx < \infty \). On this space we can define the \( L_1 \) metric to be \( d_{L_1}(g_1, g_2) = \int_{\mathbb{R}} |g_1(x) - g_2(x)| \, dx \). The \( L_1 \) metric is a continuous version of the \( \ell_1 \) metric. It is easy to check that for finite sets the notion of an \( L_1 \) metric and an \( \ell_1 \) metric coincide.

**Theorem 6.** Let \( f \) be a concave non-decreasing function with \( f(0) = 0 \) and let \( X \subset \mathbb{R} \) be finite, then \((X, d_f)\) is an \( \ell_1 \) metric space.

**Proof:** Let \( X = \{x_1, \ldots, x_n\} \). Consider \( Y = \{|x_i - x_j| : 1 \leq i, j \leq n\} \) the set of possible distances in \( X \), and order the elements of \( Y \) as \( 0 = y_0 < y_1 < y_2 < \ldots < y_m \). Let \( g \) be the piecewise linear function which

(i) coincides with \( f \) on \( Y \)

(ii) is linear on all intervals \([y_i, y_{i+1}]\) and

(iii) is constant on \([y_m, \infty)\) (that is, gets the value \( f(y_m) \) there).

Obviously \((X, d_f) = (X, d_g)\) since \( f = g \) on the set of all relevant values \( Y \), so it is enough to prove the claim for \( g \), which is also non-decreasing and concave. We now define functions \( h_{s,t} \) as follows.

\[
h_{s,t}(x) = sx \text{ if } x < t \text{ and } h_{s,t}(x) = st \text{ otherwise.}
\]

We also let \( s_i = \frac{g(y_i) - g(y_{i-1})}{y_i - y_{i-1}} \) be the sequence of slopes of \( g \). We claim that \( g \) is a convex (in particular, non-negative) combination of functions of the form \( h_{s,t} \).

The proof proceeds by induction on \( m \). If \( m = 0 \) then \( g = h_{1,0} = 0 \). For \( m > 0 \), look at the function \( \tilde{g} = g - h_{s_m,y_m} \). It is not hard to see that \( \tilde{g}(0) = 0 \), \( \tilde{g} \) is constant beyond \( y_{m-1} \) and is piecewise linear with breakpoints \( y_1, \ldots, y_{m-1} \). A piecewise linear function is concave and non-decreasing if and only if its slopes are non-increasing and non-negative, and so \( s_1 \geq s_2 \geq \ldots \geq s_m \geq 0 \), and similarly \( s_1 - s_m \geq s_2 - s_m \geq \ldots \geq s_{m-1} - s_m \geq 0 \). But, these are the slopes of \( \tilde{g} \) and it is therefore concave and non-decreasing. We may now apply the induction hypothesis to \( \tilde{g} \) and this proves the claim.

Since a sum of \( \ell_1 \)-metrics is also an \( \ell_1 \)-metric, we are left with the task of showing that for a function \( h_{s,y} \), the resulting metric \( d_h \) is an \( \ell_1 \)-metric. Notice that \( d_h(x_i, x_j) = s \cdot \min\{|x_i - x_j|, y\} \). Let \( f_i = \frac{s}{2} \chi_{[x_i, x_i+y]} \) be the function whose value is \( \frac{s}{2} \) on the interval \([x_i, x_i+y]\) and zero otherwise. It
is easy to see that for any \( x_i, x_j \in \mathbb{R} \)

\[
d_h(x_i, x_j) = s \cdot \min\{|x_i - x_j|, y\} = \int_{\mathbb{R}} |f_i(x) - f_j(x)| \, dx
\]

This shows that \( d_h \) is an \( L_1 \) metric and hence \( \ell_1 \). q.e.d.

The original motivation for studying the results of this chapter was the development of an automated software tool for reconfiguring data in a disk array. The IRM estimate was used as part of an estimate on disk utilization. This in turn was used to estimate the impact of moving some data from one disk to another. If the estimate suggested that the utilization of the disks dropped substantially then the data was moved. In such applications which use incomplete information it is imperative to have conservative estimates, otherwise major mistakes can occur. The results of the present chapter, which is based on [19], showed that the IRM estimate is safe to use. The software tool has been commercially available for more than a decade in many systems and indeed the algorithm makes very few errors in judgment. For a description of the application and one success story, see [10].

Notes for chapter 1

(1) In the language of graph theory a sequence of requests with all multiplicities \( a_i = 1 \) is a Hamiltonian cycle in the complete graph with \( n \) vertices. The seek times \( d_{i,j} \) serve as weights on the graph edges. The quantity \( W(d) = W(d, a_i) \), is the largest weight of a Hamiltonian cycle. Finding a largest weight Hamiltonian cycle is a well known problem of theoretical computer science known as MAX-TSP (maximal traveling salesman problem). For general values of \( d_{i,j} \), or even for a metric, the problem of computing \( W(d, a_i) \) is difficult, or in more technical terms, NP-Complete, see [77] for a comprehensive and readable account of the theory of NP-Completeness. However, a polynomial time algorithm of Hassin and Rubinstein, see [93], computes a Hamiltonian cycle whose weight is at least \((7/8)W(d)\). When we do not require that \( d_{i,j} = d_{j,i} \), there are sophisticated algorithms that compute cycles, of weight, at least \((35/44)W(d)\), see [104, 118]. Our formula \( E(T)(d) \), does not present an explicit cycle and is only guaranteed to be at least \((1/2)W(d)\), however, it is much simpler to compute and is universal.

(2) Given a function \( f \) we can define a semi-metric on the real line by setting \( d_{x,y} = f(|x - y|) \). We may ask if the real line with the square of this metric can be embedded in Euclidean space in a metric preserving manner. Since the real line is infinite we may need an infinite dimensional space to map it to. We therefore replace Euclidean space by Hilbert space, \( H \), the space of all sequences \( (x_1, x_2, \ldots, x_n, \ldots) \) for which \( \sum_{i=1}^{\infty} x_i^2 < \infty \). This space is equipped
with the norm \( ||(x_i)|| = \sqrt{\sum_{i=1}^{\infty} x_i^2} \). The norm, in turn, induces a metric on \( H \) given by \( d_{(x_i),(y_i)} = ||(x_i - y_i)|| \). The space \( \mathbf{R}^n \) with the Euclidean metric sits inside \( H \) as the set of all sequences of the form \((x_1, x_2, ..., x_n, 0, 0, ...). \) The question now becomes, whether we can find a map \( \phi \) from the real line to \( H \), such that \( f^2(|x - y|) = d_{\phi(x),\phi(y)} \). It is obvious that if such a map exists then any finite set of the real line \( r_1, ..., r_n \) with the induced metric is square Euclidean. A simple theorem of Menger states that the converse is also true. If any finite subset of the real line can be embedded in Euclidean space, preserving distances, then the whole line can be embedded in Hilbert space in a distance preserving manner.

In their paper, [180] Schoenberg and Von Neumann, gave an if and only if criterion for functions \( f \) such that the real line \( \mathbf{R} \) with the semi-metric \( d_f = f^2(|x - y|) \) embeds into Hilbert space. They called such functions, \( f \), screw functions. To state their criterion we allow the case \( f(x) = 0 \), for \( x \neq 0 \), which means that the map \( \phi \) does not have to be 1-1. Before stating the theorem we recall that a non-decreasing function \( \gamma(t) \), \( t \geq 0 \), gives rise to a non-negative measure on the non-negative reals. The measure of the interval \([a,b]\) is defined to be \( \gamma(b) - \gamma(a) \). The measure of a point \( a \) is defined to be \( \lim_{\varepsilon \to 0} \gamma(a) - \gamma(a - \varepsilon) \). We can use this measure to integrate functions defined on the non-negative numbers. We will denote the measure by \( d\gamma(t) \).

**Theorem 7.** A non-negative, \( f(t) \), satisfying \( f(0) = 0 \), is a screw function if and only if \( f \) has the form

\[
 f^2(t) = ct + \int_0^\infty \frac{\sin^2(ut)}{u^2} d\gamma(u)
\]

for some non-decreasing function \( \gamma(u) \) such that \( \int_0^\infty \frac{d\gamma(u)}{u^2} < \infty \).

We will not prove this result, however we can present some motivation for it. The real line is contained in Hilbert space and we can use the map \( g(t) = ct \) to see that functions of the form \( f^2(t) = ct \) are screw functions.

We note that the set of square Euclidean semi-metrics on a space \( X \) is closed under the operation of taking non-negative linear combination, hence, if \( f_1 \) and \( f_2 \) are screw functions then so is \( f = c_1f_1 + c_2f_2 \), where \( c_1, c_2 \geq 0 \). To see this, we note that if for \( X \), a finite subset of the real line \( \phi_1 \) and \( \phi_2 \) are mappings of \( X \) to Euclidean spaces of dimensions \( m_1 \) and \( m_2 \), which are distance preserving for \( d_{f_1} \) and \( d_{f_2} \) respectively, then by Pythagoras' theorem, the map which sends \( x \in X \) to \((\sqrt{c_1}\phi_1(x), \sqrt{c_2}\phi_2(x))\) in \( m_1 + m_2 \) dimensional space is distance preserving for \( d_f \).
Another source for screw functions are the exponential maps \( \phi_s(t) = e^{ist} \), where we identify the complex numbers with \( \mathbb{R}^2 \).

In representation theoretic terms, these are the one dimensional unitary representations of the real line, the characters. It is a simple exercise in trigonometry to verify that the mapping \( t \rightarrow (1/2)(\cos(2ut), \sin(2ut)) \), proves that \( f(t) = |\sin(ut)| \) is a screw function. We see that any function of the form \( f^2(t) = c_0t + \sum_{i=1}^k c_isin^2(u_it) \) is a screw function.

Such sums correspond to integration against a sum of Dirac measures, which are precisely the measures corresponding to finite step functions \( \gamma \) in the theorem. More precisely, let \( \gamma = c_0\chi_{[0,\infty)} + \sum_{i=1}^k c_iu_i^2\chi_{[u_i,\infty)} \), where \( \chi \) denotes the characteristic function of an interval, then by definition \( f^2 = \int_0^\infty \frac{\sin^2(ut)}{u^2}d\gamma(u) \). By point-wise approximations of non-decreasing functions by a sequence of finite step functions we see that all the functions in the theorem are indeed screw functions. Point-wise convergence is sufficient because of Menger's theorem. The hard part is showing the converse, that these are all the functions. We refer the reader to, [180], or the book [181] for the proof.

To show that \( f_\alpha(x) = x^\alpha \) is a screw function for all \( 0 < \alpha < 1 \), the corresponding integral representation is given by the identity \( t^{2\alpha} = c_\alpha \int_0^\infty u^{-(1+2\alpha)}\sin^2(ut)du \), for a positive constant \( c_\alpha \).

Von-Neumann and Schoenberg proved more generally an analogous characterization for screw functions of \( \mathbb{R}^m \), namely functions \( f \) for which \( \mathbb{R}^m \) equipped with the metric \( d_f = f(\|x-y\|) \), is square Euclidean.

We let \( S^{m-1} \) denote the unit sphere in \( \mathbb{R}^m \) and let \( \sigma_{m-1} \) denote the normalized rotationally invariant Haar measure on \( S^{m-1} \). We can define a function

\[
\Omega_m(t) = \int_{S^{m-1}} e^{i(x,y)}d\sigma_{m-1}(y)
\]

where \((x,y)\) denotes the inner product and \( x \in \mathbb{R}^m \) is any vector with \( \|x\| = t \). Computing in polar coordinates with the vector \((t,0,\ldots,0)\) one obtains \( \Omega_m(t) = \frac{\int_0^\pi e^{i\theta} \sin^{m-2}(\theta) d\theta}{\int_0^\pi \sin^{m-2}(\theta) d\theta} \). From this expression one can also obtain a power series expression

\[
\Omega_m(t) = \sum_{j=0}^\infty (-1)^j \frac{1}{2j!(m+2)\ldots(m+2j-2)} t^{2j}
\]

The function \( \Omega_m \) is strongly related to the order \( m \) Bessel function. In addition, the expansion shows that \( \Omega_1(t) = \cos(t) \). The theorem for general \( m \) states that \( f \) is a screw function of \( \mathbb{R}^m \) if and only if

\[
f^2(t) = \int_0^\infty \frac{1 - \Omega_m(tu)}{u^2}d\gamma(u)
\]
for a non-decreasing function $\gamma$ satisfying
\[
\int_0^\infty \frac{1}{u^2} d\gamma(u) < \infty
\]

For a very clear account of this and many other related results, we refer the reader to [181] chapters 1 and 2.

(3) The sequence of containments in lemma 2 can be extended to include $\ell_2 \subset \ell_1$. To see this, we may consider half-spaces, namely, sets of the form $w \cdot x \geq c$, for a fixed vector $w$ and some constant $c$. Given a finite set of points $x_i \in \mathbb{R}^m$, we consider the set of half spaces which separate at least two of the points in the set. There is a translation and rotation invariant measure $d\mu$ on the set of all half-spaces and the measure is finite on the set of separating half-spaces. To each point we assign the subset of half-spaces which contain it. The measure of the set of half-spaces which separate a particular pair of points $x_i$ and $x_j$ is proportional to their $\ell_2$ distance, showing that this map to $\ell_1$ after a suitable scaling is distance invariant.

(4) The difference between $\ell_1$ and $\ell_2^2$ embeddability turns out to reflect the difference between classical mechanics and quantum mechanics. One thinks of measurements which are performed simultaneously in two different places. Assume that the set of results of experimental measurements performed in one lab are given by some binary random variables $A_1, \ldots, A_k$ and that the experiments in the other lab yield binary random variables $B_1, \ldots, B_l$. We define a binary correlation random variable $A_i \Delta B_j$ to have value 1 if $A_i = B_j$ and zero otherwise. For a random variable $X$ we denote its expectation by $E(X)$. We say that a vector is a $k,l$ bipartite correlation vector if it is the concatenation of the vectors $(E(A_i))$, $(E(B_j))$ and $(E(A_i \Delta B_j))$, for $1 \leq i \leq k$ and $1 \leq j \leq l$. In short the vector is a possible vector of expectations of experiments and their pairwise inter-lab correlations. Correlation vectors are related to $\ell_1$ metrics in the following way. A vector is a correlation vector if and only if we can find vectors $v_i, w_j \in \mathbb{R}^n$, such that $\|v_i\|_1 = E(A_i)$, $\|w_j\|_1 = E(B_j)$ and $\|v_i - w_j\|_1 = E(A_i \Delta B_j)$. To see why this result holds, we consider the extreme case where all the random variables are set deterministically to either the value 0 or the value 1. We notice that in this case the required distances correspond precisely to the cut metric $d_S$, where $S$ is the set of random variables which are set to 0. All other joint distributions of the binary random variables are simply convex combinations of these extremal settings. The set of all possible correlation vectors is a polytope. The faces of the polytope are given by linear inequalities that the correlation vectors must satisfy. These inequalities are called Bell
inequalities. When \( k = l = 2 \), there is one non obvious Bell inequality, the CHSH inequality which states that

\[
E(A_1 \Delta B_1) - E(A_1 \Delta B_2) - E(A_2 \Delta B_1) - E(A_2 \Delta B_2) \leq 0
\]

The inequality can easily be verified by checking it for all 16 possible 0/1 deterministic settings for \( A_1, A_2, B_1, B_2 \).

We notice that the main difference between the notion of \( l_1 \) that was just used and the notion of an \( l_1 \) metric which we have used previously is that we require only some of the distances between elements, namely, those between the points \( v_i \) and the points \( w_j \) to be \( l_1 \), while stating nothing about internal distances between the \( v_i \) or between the \( w_j \). We have also implicitly added one more random variable \( C \), which is deterministically zero. We always map it to the origin and we insist that distances from \( C \) are preserved, which leads to the conditions \( ||v_i||_1 = E(A_i) \), \( ||w_j||_1 = E(B_j) \). We can summarize all this by considering the graph \( G = G_{k,l} \) on vertices \( A_i, B_j, C \), with edges, \( (A_i, B_j) \), \( (A_i, C) \), \( (B_j, C) \), for all \( 1 \leq i \leq k \) and \( 1 \leq j \leq l \). We say that a function on the edges is in \( l_1(G) \) if the vertices can be mapped to \( l_1 \) so that the weights equal the \( l_1 \) distances of the images. We can use the same procedure to define \( l_p(G) \) or \( l_2^2(G) \) for any graph \( G \).

The set of results of a classical experiment corresponds to a set of weights in \( l_1(G_{k,l}) \).

Now what about the results of quantum mechanical experiments? If we restrict ourselves to binary random variables \( A_i, B_j \) which are balanced, namely \( E(A_i) = Pr(A = 0) = E(B_j) = Pr(B = 0) = 1/2 \) then a remarkable result of Tsirelson, \([176, 177]\), states that a vector is a quantum correlation vector if and only if it is in \( l_2^2(G_{k,l}) \). To explain the connection between \( l_2^2 \) and quantum mechanics we would need to explain the basic set-up of the theory, which would be a bit too long for this note. Instead we refer the reader to \([14]\), from which this material is taken, where a sketch of the proof can be found along with other results.

(5) We have seen that within all metric spaces there is a sequence of containments \( \ell_1 \subset \text{Hypermetrics} \subset \ell_2^2 \). Consider the set of invariant metrics on \( \mathbb{R}^m \), i.e., metrics of the form \( d_f = f(||x - y||) \), for some function \( f \). We denote this set by \( \text{Imetrics}(m) \). We let \( \ell_1(m), \text{Hypermetrics}(m), \ell_2^2(m) \), be the subset of \( \text{Imetrics}(m) \) with the appropriate property. Let \( \text{Concave}(m) \) be the set of invariant metrics with \( f \) being concave. We can summarize our results by the sequence of containments

\[
\text{concave}(1) \subset \ell_1(1) \subset \text{Hypermetrics}(1) \subset \ell_2^2(1) \subset \text{Imetrics}(1)
\]
However, it is not clear, which of these containments are proper, apart from the trivial fact that there are invariant metrics, which do not come from a concave function.

We may also ask for a clarification of the hierarchy for $m > 1$. In that case the work of Von-Neumann and Schoenberg, [180] shows that not all concave invariant metrics are $\ell^2_2$, since the $\ell^2_2$ condition imposes some differentiability conditions on $f$, while $H_{s,t}$ are not differentiable. On the other hand the functions $f(x) = x^c$, $0 < c < 2$ are in $\ell^2_2$ for all $n$. 
CHAPTER 2

A fractal model for storage system activity

In the first chapter we introduced a very simple model, the IRM model, for storage system activity. One aspect that we did not fully consider is the dynamics of the system, how activity evolves over time. There was no specific aspect of time in the model. All the parameters of the model were estimated directly from observations and any dynamics was again based on measurements in different time periods, the time periods typically being rather long, on the order of minutes. Sometimes we want to capture phenomenon on a much finer time scale.

If one observes a trace from a real storage system, it is likely that several features will be observed. One typical observation is that I/Os tend to be very bursty. There may be long periods of little activity followed by bursts of great activity over short time spans. This is also true of addresses. Some locations are far more popular (burst with activity) than others. Another feature is that these bursts of activity in time and location (space) are correlated. The popular locations can vary in time. Such typical behavior was captured recently using an interesting new family of models for traces, which are based on fractal measures on the unit square. These models were introduced by Wang, Ailamaki and Faloutsos, [183], while an earlier version, which considers only the timing of requests, was introduced in [182]. For reasons that will be clearer later, the models were called PQRS models, but we will prefer the name bias models. The (parameter dependent) models in the family have the property that some combinations of requested addresses and access times are far more likely than others. This property is responsible for the bursty, space-time localized, nature of the I/O access patterns produced by the models and the strong spatio-temporal dependencies that the pattern displays, just like real I/O access patterns. Another great feature of these models, which is related to their fractal nature, is that they require very few parameters.

The work on PQRS/bias models which was presented in [183] was based on simulations and empirical observations, rather than mathematical analysis. In order to calculate hit ratios and average seek distances the authors simulated the model. The results led to some interesting observations. They observed, for example, that when space-time correlations are ignored average seek distances in the models are too high and hit ratios too low. They also noted that the entropies associated with the model play an important
role, both in setting up model parameters and in the behavior of the access pattern produced by the model.

In this chapter we will introduce the model and analyze it mathematically, proving the basic observations from the paper. The analysis will, among other things, naturally lead us to some basic mathematical results such as Stirling’s formula and the central limit theorem which are generally useful in performance analysis. Once we know how to compute things we do not need to carry out simulations, whose results can be calculated in advance.

1. Definition of the model

Recall that a trace of I/O activity is a sequence of pairs \((s(i), t(i))\), \(i = 1, \ldots, N\), where \(s(i)\) is a storage address and \(t(i)\) is a time stamp. The pair \((s(i), t(i))\) represents an I/O request to address \(s(i)\) which was requested at time \(t(i)\). Let \(T\) be a time interval during which we want a trace to be active, i.e., \(t(i) \in T\). Let \(S\) denote the range of storage addresses for the I/O in the trace, i.e., \(s(i) \in S\).

Let \(q \geq 2\) be some integer which will serve as the basis for expanding addresses and times, in other words, \(0, \ldots, q - 1\) will serve as digits. Usually, in practice, one considers \(q = 2\) and the "digits" are simply the bits which compose the addresses and time stamps. We assume for the sake of simplicity that the size of the address range \(|S|\) and the clock time range \(|T|\) are given by \(0, \ldots, q^h - 1\), for some \(h\).

It is mathematically convenient to realize the address space via subsets of the unit interval. We let \(I^S = [0, 1]\). For \(i = 0, \ldots, q^h - 1\) we define the \(i\)'th \(q\)-ary interval of level \(h\) to be the sub-interval \(I_{i,q,h}^S = \left[ \frac{i}{q^h}, \frac{i+1}{q^h} \right] \). The sub-interval \(I_i^S\) will represent the address \(i\). We realize clock times in the same manner, defining \(I^T = [0, 1]\) and \(I_{i,q,h}^T = \left[ \frac{i}{q^h}, \frac{i+1}{q^h} \right] \). Consequently, a trace entry given by an address \(s(i)\) at time \(t(i)\) is represented by the sub-square \(I_{s(i)}^S \times I_{t(i)}^T\) in the square \(I^S \times I^T\) which represents all possible address locations in \(S\) over the entire time interval \(T\).

Consider an integer \(m\) in the range \(0 \leq m < q^h\). We may write \(m\) as a \(q\)-ary sequence of length \(h\), \(m = m_0, m_1, \ldots, m_{h-1}\). Conversely we identify a \(q\)-ary sequence of length \(h\) with the number \(m = \sum_{l=0}^{h-1} i_l q^l\). We will think of a \(q\)-ary sequence \(m\) of length \(h\) either as a sequence of length \(h\) or as an integer in the range \([0, q^h - 1]\) interchangeably without further mention.

We will think of a trace entry \(s(i), t(i)\) where \(s(i) = s(i)_0 s(i)_1 \ldots s(i)_{h-1}\), \(t(i) = t(i)_0 \ldots t(i)_{h-1}\) as a \(q^2\)-ary sequence of length \(h\), with \(q^2\)-ary digits \(s(i)_0, t(i)_0\) up to \(s(i)_{h-1} t(i)_{h-1}\). Thus we have paired the most significant digits of the address and clock time into a single most significant address-clock digit, repeating for the other digits as well.

Consider a probability measure \(\nu\) on \(q^2\)-ary digits, namely, on pairs \(0 \leq k, l < q\). For any \(h\), \(\nu\) induces a measure \(\mu = \nu^h\) on trace entries, the probability of a trace entry \((s(i), t(i))\), or of the corresponding sub-square
1. Definition of the model

A q-ary bias model of level \( h \) is given by \( h, q^2 \)-ary probability measures \( \nu_0, \ldots, \nu_{h-1} \). The probability of a single trace entry is given by

\[
\mu(s(i), t(i)) = \prod_{j=0}^{h-1} \nu_j(s(i)_j, t(i)_j)
\]

A trace of length \( a \) is generated by sampling \( \mu \), independently, \( a \) times and re-arranging according to the time stamp.

A q-ary bias model is said to be self-similar if all the \( \nu_i \) are the same, i.e., \( \mu = \nu^h \).

In the paper [183] where such models were first introduced \( q \) was chosen to be two and the bias models were self similar. The 4 probabilities of the \( 2^2 \)-ary digits where denoted by \( p, q, r, s \), hence the name PQRS models.

We can also take \( h = \infty \) which simply means that we assign a measure to a sub-square of any level \( h \). We consider the map \( T_q \) from \( I^S \times I^T = [0, 1] \times [0, 1] \) to itself, which maps a point \((x, y)\) to the point \((qx, qy)\). In terms of the \( q \)-ary digit expansion of \( x \) and \( y \), \( T_q \) is simply a shift. We notice that in the self-similar case, the measure \( \mu_\nu \) on \( I^S \times I^T \) is invariant with respect to \( T_q \), namely \( \mu_\nu(T_q^{-1}(A)) = \mu_\nu(A) \), this is easily seen from the shift description.

1.1. Marginal spatial and temporal measures. A bias measure \( \mu = \mu(\nu_0, \ldots, \nu_{h-1}) \) induces spatial and temporal marginal measures \( \mu^S \) and \( \mu^T \) on integers in the range \([0, q^h - 1]\) via integration (summation). Stated otherwise, \( \mu^T(j) \) is the probability that time \( j \) will be the clock time of an I/O in the bias model. Specifically

**Definition 14.** Given a bias measure \( \mu = \mu(\nu_0, \ldots, \nu_{h-1}) \), we define the **temporal marginal measure** \( \mu^T \) to be the measure on \( I^T \) that assigns to the sub-interval \( I^T_{j,q,h} \) the probability

\[
\mu^T(j) = \mu^T(I^T_{j,q,h}) = \sum_{i=0}^{q^h-1} \mu(i,j)
\]

The definition of \( \mu^S \), the **spatial marginal model**, is similar for addresses.

The time marginal \( \mu^T \) can be described as follows. We notice from the description of \( \nu_k \) that the probability that a given clock time digit \( j_k \), \( k = 0, \ldots, h - 1 \), in any I/O will have the value \( l \) is \( \nu^T_k(l) = \sum_{m=0}^{q-1} \nu_k(l, m) \). Since the digits are drawn independently the marginal distribution \( \mu^T \) is given by \( \mu^T(j) = \prod_{k=0}^{h-1} \nu^T_k(j_k) \), which is simply a product of its digits.
2. A Fractal Model for Storage System Activity

Definition 15. We say that a bias model \( \mu \) is an \( I \)-model if \( \mu = \mu^S \times \mu^T \). This is the same as saying that \( \nu_k = \nu^S_k \times \nu^T_k \) for \( 0 \leq k < h \).

2. Average seek distance calculations

Having set up the I/O model we would like to calculate some related quantities which are relevant to the performance of the storage system. We first consider the average seek distance between I/O requests. It will be convenient to normalize seek distances so that the maximal seek distance in the disk is 1. This coincides with the distance in the unit interval. The distance between two level \( h \) addresses \( i_1 \) and \( i_2 \) will be \( |i_1 - i_2|/q^h \).

Theorem 8. Consider a \( q \)-ary self similar bias model with bias measure \( \nu = \nu(m, n) \).

1) Assume \( h = \infty \). Let
\[
A = \sum_{m=0}^{q-1} \frac{\sum_{i,j=0}^{q-1} \nu(m,i)\nu(m,j)|i-j|}{(\sum_{l=0}^{q-1} \nu(m,l))^2}
\]
and
\[
B = \sum_{m=0}^{q-1} \frac{\sum_{i,j=0}^{q-1} \nu(m,i)\nu(m,j)}{(\sum_{l=0}^{q-1} \nu(m,l))^2}
\]
then the average seek distance in the model is
\[
E(S_\nu) = \frac{A}{q-1 + B}
\]

If the level \( h \) is finite then the average seek distance \( E(S_\nu,h) \) satisfies
\[
|E(S_\nu,h) - E(S_\nu)| \leq q^{-(h+1)}
\]

2) Among all \( q \)-ary bias models which share fixed spatial and temporal marginal measures \( \mu^T \) and \( \mu^S \) the average seek distance (seek time) is maximized by the \( I \)-model \( \mu^T \times \mu^S \). The average seek distance or seek time of an \( I \)-model depends only on the spatial marginal \( \mu^S \).

Proof: At any given time \( t \) we may compute the average seek between a request \( R_1 \) and the following (time-wise) request \( R_2 \) inductively. Consider first a request which fell time-wise in the interval \( J^T_m = J^T_{m,q,1} = [\frac{m}{q}, \frac{m+1}{q}] \). The next request time-wise will also fall within this time interval with probability approaching 1 as the total number of requests tends to infinity. Given that the time of the requests \( R_1 \) and \( R_2 \) is in \( J^T_{m,q} \), the probability that their addresses were in different spatial intervals \( J^S_j = J^S_{j,q,1} = [\frac{j}{q}, \frac{j+1}{q}] \) and \( J^S_i = J^S_{i,q,1} = [\frac{i}{q}, \frac{i+1}{q}] \) is given by
\[
\frac{\nu(m,i)\nu(m,j)}{(\sum_{l=0}^{q-1} \nu(m,l))^2}
\]
Let \( F \) be a probability distribution on an interval \( K = [0, a] \). Let \( b \geq a \). Consider the shifted distribution \( F_b(x) = F(x-b) \) on the interval \( J = [b, a+b] \). Let \( S \) denote the average distance between a point in \( K \) and a point in

...
2. AVERAGE SEEK DISTANCE CALCULATIONS

We have \[ S = \int_a^b \int_{a+1}^{b+1} (x-y) dF_b(x)dF(y) = \int_a^b xdF_b(x) - \int_b^{b+1} ydF(y) = b \] Since the distribution is shifted. Since the probability distributions on \( J_{j,q,1}^S \) and \( J_{i,q,1}^S \) are shifted in a bias model the average distance between the requests is \( |i-j|/q \).

Summing over all \( i, j \) and \( m \) and taking into account the probability that \( R_1, R_2 \) are in the time interval \( J_m^T \) we obtain that the contribution to the average seek distance coming from pairs of requests which fell into different spatial intervals is given by

\[
A/q = \sum_m \left( \sum_{i \neq j} \frac{\nu(m,i)\nu(m,j)|i-j|/q}{(\sum_k \nu(m,k))^2} \right)
\]

and that the probability of having consecutive requests in different spatial sub-intervals is

\[
B = \sum_m \left( \sum_{i \neq j} \frac{\nu(m,i)\nu(m,j)}{(\sum_k \nu(m,k))^2} \right)
\]

This is not surprising, since at the level of the sets \( J_i^T, J_i^S \) a bias model is the same as an IRM with different spatial probabilities at the different time intervals \( J_i^T \). The difference from the IRM model happens when we consider requests that fell into the same spatial interval. In the IRM model they would not contribute. By self similarity of the model, as represented by the mapping \( T_q \) on the spatial location, the average distance between requests with the same first spatial digit is the average seek distance between requests divided by \( q \). We conclude that the average seek distance \( E(S_{\nu}) \) satisfies the recursion

\[
E(S_{\nu}) = A/q + (1 - B)E(S_{\nu})/q
\]

from which the formula in part (1) of the theorem easily follows.

The models with bias measure \( \nu \) and levels \( h \) and infinity coincide on the first \( h \) address bits. Thus the differences are restricted to spatial intervals of size at most \( q^{-(h+1)} \) completing the statements in part (1).

To prove the second part, we note that by continuity of the seek time function, if \( E_{st}(\nu, h) \) denotes the average seek time in the bias model with bias measure \( \nu \) and level \( h \) then \( \lim_{h \to \infty} E_{st}(\nu, h) = E_{st}(\nu, \infty) \). When \( h \) is finite, the bias model is a dynamic IRM model, with different probabilities for the different time intervals \( \mu_{i,q,h} \). The only difference from what we considered in the first chapter is that we consider the seek time function as being continuous. Correspondingly, the average seek time expression is

\[
(13) \quad \sum_i \mu_i^T \left( \int_0^1 \int_0^1 F(|u-v|) d\mu_i^S(v) d\mu_i^S(u) \right)
\]

where \( d\mu_i \) is the spatial measure during the time interval \( I_{i,q,h} \). For a given \( n \geq h \) consider the grid of points of the form \( (k/2^n, l/2^n), 0 \leq k, l \leq 2^n - 1 \). Consider the Riemann sums for the integrals in (13), with respect to this grid sub-division. The resulting expression for each integral is easily seen to be the average seek time in an IRM model which depends on the time
interval $I_{i,q,h}^T$. If the measure $\mu$ is an $I$-bias measure then all the integrals (Riemann sums) are the same and equal to the average seek time assuming a static model whose spatial distribution is the spatial marginal $\mu^S$. The statement that this average seek time is maximal is just a special case of theorem 4 on the super-additivity of the IRM seek time estimate. Letting $n \to \infty$ gives us the required result in the limit. \textit{q.e.d.}

We make a couple of remarks on the theorem. The formula in part (1) can be easily generalized to the non self-similar case where the bias measures $\nu_j$ on the digits depend on the level. However, the average seek distance will be given by an infinite series sum rather than a finite expression. Part (2) also extends to this case since we never used the assumption that the measure is level independent.

Consider the binary bias model with $\nu(0,0) = \nu(1,0) = \nu(0,1) = \nu(1,1)$. This bias model corresponds to requests with a uniformly random spatial location at uniformly random times. In this case $A + B = 1/2$ and we get an average seek of $E(S_\nu) = 1/3$. But in this model an average seek is simply given by the average distance between two points in the unit interval. This distance can be calculated as

$$\int_0^1 \int_0^1 |y - x| \, dx \, dy = 2 \int_0^1 \int_x^1 (y - x) \, dx \, dy = 1/3$$

so we got a self-similar (fractal) argument for computing this integral.

### 3. Cache hit ratios and entropy

In most storage systems the bulk of the data is stored in high capacity, cheap and slow storage devices, such as disk drives. However, there is usually a much smaller amount of expensive, fast storage, known as a cache in which we try to hold the most popular data items at any given moment. If at the time of a read request, the requested item happened to be in the cache, we do not need to go to the slow storage, instead we simply transfer the item from the cache to the user. If the request is serviced from the cache we say that the request was a (read) hit.

**Definition 16.** The ratio of the number of requested data items which were served from the cache (fast response), to the total number of data requests is known as the hit ratio.

Basically, we try, at any given moment, to have in the cache the data items that are most likely to be requested so that the hit ratio will be as large as possible, leading to faster service.

**Definition 17.** A strategy that determines the content of the cache at any given moment will be called a caching algorithm.

There are usually limitations on the amount of data that we can push in and out of cache in a given time frame. Frequently, there are also limitations
on the amount of computations that we can perform in order to make a decision on the cache content. Thus, in reality, caching algorithms tend to be simple to plan and to execute and do not always optimize the expected hit ratio.

The efficiency of a caching strategy will depend in general on the user access pattern (the request trace), the size of the cache and on the caching algorithm. In this section we will assume that the trace is generated by a bias model. We will consider two caching algorithms. One is optimal given the bias model setting. The other will be sub-optimal, but will be the simplest possible algorithm to manage. We will see that the hit ratios for the two algorithms are strongly related to the concept of entropy so we begin with a few basic facts about entropy.

3.1. Basic properties of entropy. We recall from the first chapter that given a probability distribution \( p_1, \ldots, p_m \geq 0, \sum_{i=1}^{m} p_i = 1 \), we may define the (Shannon) entropy of the distribution as

\[
H(p_1, \ldots, p_m) = -\sum_{i=1}^{m} p_i \log_2(p_i).
\]

**Definition 18.** Given a pair \( X,Y \) of random variables with a finite set of values, the *joint entropy* \( H(X,Y) \), of \( X,Y \) is the entropy of their joint distribution \( p(x,y) \).

\[
H(X,Y) = -\sum_{x,y} p(x,y) \log(p(x,y))
\]

We also define \( H(Y|X) \), the *conditional entropy* of \( Y \) conditioned on \( X \) as

\[
H(Y|X) = -\sum_{x,y} p(x,y) \log(p(y|x))
\]

The following basic lemma relates the two concepts.

**Lemma 3.** The identity

\[
H(X,Y) = H(X) + H(Y|X)
\]

holds. As a consequence, if \( X,Y \) are independent random variables then

\[
H(X,Y) = H(X) + H(Y)
\]

**Proof:** We compute

\[
H(X,Y) = -\sum_{x,y} p(x,y) \log(p(x,y)) = -\sum_{x,y} p(x,y) \log(p(x)p(y|x))
\]

\[
= -\sum_{x,y} p(x,y) \log(p(x)) - \sum_{x,y} p(x,y) \log(p(y|x)) = H(X) + H(Y|X)
\]
as required. If \( X,Y \) are independent random variables then by definition \( p(x,y) = p(x)p(y) \) or equivalently \( p(y|x) = p(y) \) and we obtain from the above computation the second claim. *q.e.d.*

We add one more notion.
Definition 19. The **mutual information** $I(X,Y)$ of a pair of finite random variables $X,Y$ is given by

$$I(X,Y) = \sum_{x,y} p(x,y) \log \left( \frac{p(x,y)}{p(x)p(y)} \right)$$

Given a pair of random variables $X,Y$ we may consider independent random variables $\tilde{X}, \tilde{Y}$ having the same distributions as $X$ and $Y$ respectively. From the definition we see that $I(X,Y) = D((X,Y) || (\tilde{X}, \tilde{Y}))$ where $D$ refers to the relative entropy which was defined in chapter 1. In particular, from the proof of proposition 1 we see that

$$I(X,Y) \geq 0$$

with equality if and only if $X$ and $Y$ are independent random variables. The following lemma relates the mutual information to the previously defined concepts.

**Lemma 4.** The following identity holds

$$(16) \quad I(X,Y) = H(X) - H(X|Y)$$

In particular $H(X) \geq H(X|Y)$.

**Proof:** We compute

$$I(X,Y) = \sum_{x,y} p(x,y) \log \left( \frac{p(x,y)}{p(x)p(y)} \right) = \sum_{x,y} p(x,y) \log \left( \frac{p(x|y)}{p(x)} \right)$$

$$= -\sum_{x,y} p(x,y) \log(p(x)) + \sum_{x,y} p(x,y) \log(p(x|y)) = H(X) - H(X|Y)$$

The inequality follows from the non-negativity of $I$. q.e.d.

Definition 20. Given a bias model $\nu, h$ we define its entropy to be the entropy of its bias measure $H = H(\nu)$. Similarly we define the **spatial entropy** and **temporal entropy** to be the entropies of the marginal distributions, $H^{S} = H(\nu^{S})$ and $H^{T} = H(\nu^{T})$.

3.2. **Static cache algorithms.**

Definition 21. A **static cache algorithm** is a cache algorithm which decides in advance (or after a while) which addresses will reside in cache. Once the decision is made nothing enters or leaves the cache.

Static cache algorithms are obviously the simplest cache algorithms to implement. Given a probability distribution $\mu^{S}$ on the addresses, it is easy to design an optimal static cache algorithm. Consider a cache of size $K$. We sort the addresses from the highest probability to the lowest, in our case say $s^{1}, s^{2}, \ldots, s^{q}$. We then place in cache the addresses which are most likely to be called, namely $s^{1}, \ldots, s^{K}$. The hit ratio for such a method is simply $\sum_{j=1}^{K} \mu^{S}(s^{j})$. This is true not only for bias models, but for any process that
generates addresses and for which such a limiting distribution \( \mu^S \) on the addresses exists.

### 3.3. Optimal cache algorithm.

We can find the optimal cache algorithm for a bias model by following the same strategy that led us to the optimal static algorithm. Assume that the size of the cache is \( K \). Consider a given time stamp \( t = t_0 t_1 \ldots t_{h-1} \). We can sort the addresses \( s \) in descending order according to the probability \( \mu(t, s) \) that an I/O with time stamp \( t \) will have address \( s \). We then place in cache at time \( t \) the \( K \) addresses with the highest probability. This strategy is not static since the ordering depends on the time stamp \( t \), in fact this strategy may require wholesale changes of the cache content at a very high frequency.

In reality we cannot implement optimal caching strategies since we do not know the probability distribution \( \mu(t, s) \).

### 4. Asymptotic hit ratio computations

We quantify the performance of the preceding cache algorithms when \( h \) becomes large. We will consider a bias model given by a \( q \)-ary probability distribution \( \nu \). For every \( h \) we consider a cache of size \( K_h = 2^{\alpha h} \) and consider the limit of the hit ratio as \( h \to \infty \) of such a system with a given cache algorithm, assuming it exists.

**Theorem 9.** For the optimal static cache algorithm the limit as \( h \to \infty \) of the hit ratio on a cache of size \( 2^{\alpha h} \) is 1 if \( \alpha > H(\nu^S) \) and is 0 if \( \alpha < H(\nu^S) \). For the optimal cache algorithm the limit is 1 if \( \alpha > H(\nu^S | \nu^T) \) and is 0 if \( \alpha < H(\nu^S | \nu^T) \). Consequently, the mutual information controls the difference in performance between the optimal, and the optimal static algorithms.

**Proof:** for simplicity we consider the case \( q = 2 \), the case of \( q > 2 \) is similar. Ignoring the time stamp, the address bits \( X_i \) in the bias model are i.i.d., with the common distribution function \( X \) with \( L = Pr(X = 0) = \mu^S(0) \) and \( R = Pr(X = 1) = \mu^S(1) \). The number of bits set to 1 in an address is the random variable \( \sum_{i=1}^h X_i \). Let \( E(X), V(X) \) and \( \sigma(X) \) denote the average, variance and standard deviation of \( X \) respectively (see the appendix for basic definitions and results). We have, \( E(X_i) = R \), \( V(X_i) = R - R^2 = RL \) and so \( \sigma(X_i) = \sqrt{RL} \). We conclude from the weak law of large numbers (see appendix) that for any \( c > 0 \), \( Pr(\left| \sum_i X_i - hE(X) \right| \geq c\sqrt{RL}h \leq \frac{1}{c^2} \). In particular as \( h \to \infty \), for any \( \delta > 0 \) we have \( Pr(\left| \sum_i X_i - hR \right| \geq \delta h) \to 0 \).

Assume that \( R < L \), the other case being treated similarly. We see that to obtain a hit ratio approaching 1, it is enough to place in cache all addresses with at most \((R + \delta)h \) bits set to 1, for any fixed \( \delta > 0 \). To obtain an estimate for the number of such addresses we assume for simplicity that \((R + \delta)h\) is an integer. For a pair of non negative integers \( n \geq m \) we let

\[
B(n, m) = \frac{n!}{m!(n-m)!}
\]
denote the corresponding binomial coefficient. The number of addresses with exactly \((R + \delta)h\) set to 1 is given by \(B(h, (R + \delta)h)\). To estimate binomial coefficients we want to estimate factorials. and that is the content of Stirling’s celebrated formula

\[
\ell! \sim \sqrt{2\pi\ell} \frac{\ell^\ell}{e^\ell}
\]

where \(\sim\) means that the ratio of both sides approaches 1 as \(\ell \to \infty\). For now we will only prove a weak version of it. We will complete the proof later on.

**Lemma 5.** (A weak version of Stirling’s formula) There exists a constant \(C > 0\), such that

\[
limit_{\ell \to \infty} \frac{\ell!}{C \sqrt{\ell} \frac{\ell^\ell}{e^\ell}}
\]

**Proof of lemma:** We have

\[
\log(\ell!) = \sum_{k=1}^{\ell} \log(k) = \sum_{k=1}^{\ell} \int_{1}^{k} \frac{1}{x} dx = \int_{1}^{\ell} \frac{\ell - [x]}{x} dx
\]

where \([x]\) is the integral part of \(x\). The last equality since \(\frac{1}{x}\) only appears in the integrals corresponding to \(k = [x] + 1, \ldots, \ell\). We have \(\ell - [x] = \ell - x + (x - [x])\). Writing \(\tilde{x} = x - [x]\) for the fractional part of \(x\) we obtain

\[
\log(\ell!) = \int_{1}^{\ell} \left(\frac{\ell - [x]}{x} - 1 + \frac{1}{2} \frac{1}{x} + \frac{\tilde{x} - 1/2}{x}\right) dx
\]

\[
= \ell \log(\ell) - \ell + 1 + (\log(\ell)/2) + \left(\int_{1}^{\ell} \frac{\tilde{x} - 1/2}{x} dx\right)
\]

We note that \(\int_{1}^{\infty} \frac{\tilde{x} - 1/2}{x} dx\) converges by Dirichlet’s criterion, since \(\frac{1}{x}\) decreases to 0 and \(\int_{1}^{\ell} (\tilde{x} - 1/2) dx\) is bounded. Letting \(D\) denote the value of the integral, we get upon exponentiation that

\[
\ell! \sim e^{1 + D} \sqrt{\ell} \frac{\ell^\ell}{e^\ell}
\]

as required q.e.d.

We continue with the proof of the theorem. Let \(\beta = R + \delta\). From Stirling’s formula we get

\[
B(h, \beta h) = \frac{h!}{(\beta h)! (1 - \beta) h!}
\]

\[
\sim C \sqrt{h} \frac{h^h}{e^h} \frac{1}{(1 - \beta)^{\beta h}} \frac{e^{(1 - \beta)h}}{C \sqrt{\beta h}} \frac{1}{(\beta h)_{\beta h}} \frac{e^{(\beta)h}}{C \sqrt{\beta h} (\beta h)}
\]

\[
= \frac{1}{C} \sqrt{\frac{1}{\beta (1 - \beta) h^\beta (1 - \beta)^{1 - \beta} h}}
\]

\[
= \frac{1}{C} \sqrt{\beta (1 - \beta) h} 2^{H(\beta)h}
\]
Having obtained an estimate on the number of length \( h \) sequences with \( m = \beta h \) bits which are 1, we note that \( B_{h,m-1}/B_{h,m} = \frac{m}{h-m+1} \sim \frac{\beta}{1-\beta} \) and similarly \( B_{h,m+k}/B_{h,m} \sim (\frac{\beta}{1-\beta})^k \). Since we assumed that \( R < L \), for \( \delta \) small enough we have \( \beta < 1 - \beta \) and hence the number of addresses with at most \( \beta h \) bits set to 1 is

\[
\sum_{\ell \leq \beta h} B_{h,\ell} \sim \frac{1 - \beta}{1 - 2\beta C} \sqrt{\frac{1}{\beta(1 - \beta)\lambda}} \beta^{H(\beta)h}
\]

Since the function \( H(u) \) is continuous we see that for \( \delta \) small enough we will get that the number of such addresses is less than \( 2^{H(R)+\epsilon}h \) as required. If we assume that \( R < L \), the addresses which the optimal static algorithm places in cache are those which have as few 1 as possible. Taking now \( \beta = R - \delta \), the same calculation shows that the \( 2^{H(R)-\epsilon} \) most popular addresses will have at most \( (R - \delta)h \) bits set to 1, for some \( \delta = \delta_\varepsilon > 0 \) and hence by the weak law of large numbers the resulting hit ratio will tend to 0 proving the theorem for the static algorithm.

We consider the optimal algorithm. Let \( j \) be a time stamp and let \( D_j, E_j \) be respectively the bits which are set to 0 and 1 in \( j \). Fix \( \delta > 0 \). By the law of large numbers, with probability approaching 1, as \( h \to \infty \) we have

\[
(\mu^T(0) - \delta)h < |D_j| < (\mu^T(0) + \delta)h
\]

In addition, for almost all the addresses as above there are

\[
(\nu(0,0) - \delta)h < k < (\nu(0,0) + \delta)h
\]

address bits set to 0 among the \( |D_j| \) bits whose time coordinate is set to 0. Similarly, for almost all the addresses as above there are

\[
(\nu(0,1) - \delta)h < l < (\nu(0,1) + \delta)h
\]

address bits set to 0 among the \( |E_j| \) bits whose time coordinate is set to 0. Thus, the number of addresses which we need to place in cache is \( B(m, l)B(h - m, k) \) with \( m \sim \nu^T(0)h \), \( l \sim Pr(\nu^S = 0|\nu^T = 0)h \) and \( k \sim Pr(\nu^S = 0|\nu^T = 1)h \). It is easy to see from the previous calculations that there are at most \( 2^{H(\nu^S|\nu^T) + \epsilon} \) such addresses and that a cache with size \( 2^{H(\nu^S|\nu^T) - \epsilon} \) will result in a hit ratio tending to zero. The proofs for \( q > 2 \) can be shown by first conditioning on whether the bits in question have value zero or not and using induction. \( q.e.d. \)

By being a bit more careful, we can gain a more detailed analysis of the algorithms. For simplicity we restrict ourselves to the static algorithm and the case \( q = 2 \). The analysis is essentially equivalent to proving the central limit theorem for binary variables. At the same time the analysis will complete the proof of Sterling’s formula. The proof is given in the appendix.
Theorem 10. Let \( f(h) \) be a function which satisfies \( f(h) = o(\sqrt{h}) \) and let \( g(h) \) be any function which satisfies \((\frac{L}{R})^f(h) < g(h) < (\frac{R}{L})^f(h)\). Let \( \phi_h(c) \) be the hit ratio of the optimal static cache algorithm assuming a level \( h \) binary bias model with space marginal given by \( \nu^S_0 = L > 1/2 \) when the cache is of size \( g(h)(\frac{L}{R})^{c/\sqrt{LR}}2^H(L)h \). Then, \( \lim_{h \to \infty} \phi_h(c) = \frac{1}{2\pi} \int_{c}^{\infty} e^{-x^2/2} dx \)

Remark: Our computations have another important consequence which explains the connection between caching, entropy, information and compression. Assume that a sequence of i.i.d. \( q \)-ary digits \( X_1, X_2, ..., X_n, ... \) is generated. We wish to transmit these digits in a compact way. We can decide to encode sequences of length \( h \) by binary sequences of fixed length \( m \). We encode the \( 2^m - 1 \) most likely sequences of length \( h \) using all the length \( m \) binary sequences except the length \( m \) sequence of ones. Any sequence which is not among the \( 2^m - 1 \) most popular is encoded by the sequence of \( m \) ones followed by the exact length \( h \) \( q \)-ary sequence (where each \( q \)-ary digit is written as a binary number). Our computations show that for any \( \varepsilon > 0 \), on average at most \( H(X)h + \varepsilon \) bits are needed to encode a length \( h \) sequence, while any encoding will need at least \( H(X)h - \varepsilon \) bits on average, hence this result is optimal. In this sense entropy provides an indication of the amount of information stored in such a sequence. For more on entropy and information we refer the reader to [57].

5. Incompatibility of \( q \)-ary and \( p \)-ary bias models

The basic idea behind bias models is that they bias the digits of addresses and time stamps of trace entries. We may ask, can a \( p \)-ary and a \( q \)-ary model be compatible? for example, can a \( p \)-ary bias model also have \( q \)-ary bias. If they are not compatible, then using caching algorithms designed for a \( p \)-ary bias model will not work well on the output of a \( q \)-ary model. An important result of B. Host, [98], which we present in this section provides an essentially negative answer. We will consider I-models with an address measure \( \nu^S \) and \( h = \infty \). We first consider what we mean by non-bias in a \( q \)-ary sequence.

Definition 22. A number \( x \in [0,1] \) is said to be \( q \)-normal if in the base \( q \) expansion of \( x \) every string of length \( k \) \( a_0, ..., a_{k-1} \), of \( q \)-ary digits appears with the limiting frequency \( q^{-k} \), i.e.,

\[
\lim_{N \to \infty} \frac{\left| \{ n \leq N \mid x_{n+1} = a_0, ..., x_{n+k} = a_{k-1} \} \right|}{N} = q^{-k}
\]

Definition 23. A series \( x(0), x(1), ... \) of numbers \( x(i) \in [0,1] \) is said to be equi-distributed if for any interval \([a,b] \subset [0,1]\) we have

\[
\lim_{N \to \infty} \frac{\left| \{ n \leq N \mid x^n \in [a,b] \} \right|}{N} = b - a
\]

The following simple lemma relates the two notions.
5. INCOMPATIBILITY OF $q$-ARY AND $p$-ARY BIAS MODELS

**Lemma 6.** A number is $q$-normal if and only if the sequence $x(n) = q^n x \mod 1$ is equi-distributed.

**Proof:** Let $a_0, \ldots, a_{k-1}$ be some sequence of $q$-ary digits and let $j = \sum_{j=0}^{k-1} a_j q^{k-j-1}$. Consider the interval $I_j = [\tfrac{j}{q^k}, \tfrac{j+1}{q^k}]$, then by definition $x(n) \in I_j$ if and only if $x(n)_1 = a_0, \ldots, x(n)_k = a_{k-1}$. But, $x(n)_i = x_{n+i-1}$, so we see that the $q$ normality is equivalent to equi-distribution w.r.t. $I_j$, for all integers $j$. On the other hand, given an interval $[a, b]$, let $i_k$ be such that $\frac{i_k}{q^k} \leq a < \frac{i_k+1}{q^k}$ and $j_k$ such that $\frac{j_k}{q^k} \leq b < \frac{j_k+1}{q^k}$. Let $c_k = \frac{i_k}{q^k}$, $d_k = \frac{i_k+1}{q^k}$, $e_k = \frac{j_k+1}{q^k}$ and $f_k = \frac{j_k}{q^k}$. Then the intervals $[c_k, d_k]$ and $[e_k, f_k]$ are unions of intervals of the form $I_j$. In addition $[e_k, f_k] \subset [a, b] \subset [c_k, d_k]$ and $b - a - \frac{2}{q} \leq f_k - e_k \leq d_k - c_k \leq b - a + \frac{2}{q}$. Applying $q$ normality to the approximating intervals yields in the limit equi-distribution for the interval $[a, b]$. q.e.d.

The following theorem of H. Weyl, [185], provides a criterion for the equi-distribution of a sequence.

**Theorem 11.** A sequence $x(n)$ is equi-distributed if for all $m > 0$ we have

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi imx(n)} = 0$$

**Proof:** Given an interval $I = [a, b]$, let $\chi_I(x)$ be the function which equals 1 when $x \in [a, b]$ and 0 otherwise. Given an $\varepsilon > 0$, we can find a continuous function $f = f_{I, \varepsilon}$ on $[0, 1]$ such that $f \leq \chi_I$ and $\int_0^1 (\chi_I(x) - f(x)) dx < \varepsilon$, simply by extrapolating linearly a function with values $f(0) = f(a) = f(b) = f(1) = 0$ and $f(a + \varepsilon/2) = f(b - \varepsilon/2) = 1$, or taking $f$ to be the zero function if $b - a \leq \varepsilon$. We can similarly produce a second continuous function $h = h_{I, \varepsilon} \geq \chi_I$ such that $\int_0^1 |h(x) - \chi_I(x)| dx < \varepsilon$.

Let $\phi$ be a polynomial in $e^{2\pi i x}, e^{-2\pi i x}$, such that $|f(x) - \phi(x)| < \varepsilon$ for all $x \in [0, 1]$ and $\int_0^1 f(x) dx = \int_0^1 \phi(x) dx$. Such a polynomial is constructed in lemma 27 in the appendix. Given an interval $I = [a, b]$ we have

$$\left| \frac{N}{\sum_{n=0}^{N-1} \chi_I(x(n))} - \frac{1}{N} \sum_{n=0}^{N-1} f(x(n)) \right|$$

Writing $\phi = a_0 + \sum_{-l<m<l, m \neq 0} a_m e^{2\pi m x}$ and using the assumption of Weyl's criterion we have for $N$ large enough

$$\left( \frac{1}{N} \sum_{n=0}^{N-1} \phi(x(n)) \right) \geq \int_0^1 f(x) dx - \varepsilon \geq \int_0^1 \chi_I(x) dx - 2\varepsilon = (b - a) - 2\varepsilon.$$
and hence

\[
\left\lfloor \frac{n \leq N \mid x(n) \in [a,b]}{N} \right\rfloor \geq b - a - 3\varepsilon
\]

A similar argument using \( h \) instead of \( f \) gives the corresponding upper bound and completes the proof. \( q.e.d. \)

We can now state Host’s result

**Theorem 12.** Let \( p, q > 1 \) be integers with no common divisor. Let \( \mu \) be a \( p \)-ary bias model with \( h = \infty \) such that for any \( i \) and \( j \),

\[
\nu_i^S(j) < a < 1
\]

for some fixed \( a < 1 \). Then, almost surely w.r.t. \( \mu \), a point \( x \) is normal in the base \( q \).

**Proof:** We first need a lemma

**Lemma 7.** Let \( m \neq 0 \) be some integer. Let \( k_n \) be the period of the periodic sequence \( mq^k \mod p^n \), \( k \in \mathbb{Z} \). Then, for \( n \) large enough we have \( k_n = cp^n \) for some constant \( c > 0 \).

**Proof:** Write \( m \) as \( p^e m' \) with \( m' \) prime to \( p \). We have \( mq^i = mq^j \mod p^n \) if and only if \( m'q^i = m'q^j \mod p^{n-e} \). It is thus sufficient to consider integers \( m \) prime to \( p \) since \( p^{n-e} = p^{-e}p^n \). It is well known that the residues modulo \( p^n \) which are prime to \( p \) form a group \((\mathbb{Z}/p^n\mathbb{Z})^*\). We conclude that if \( m \) is prime to \( p \) we have \( mq^i = mq^j \mod p^n \) if and only if \( q^{i-j} = 1 \mod p \) by cancellation of the mutual term \( mq^j \). Thus, the period is given by the order of \( q \) in the group of prime to \( p \) residues. Let \( k_2 \) be as in the statement of the theorem, the minimal period of the sequence \( mq^k \mod p^2 \). We have seen that this is the same as the minimal integer \( l \) for which \( q^l = 1 \mod p^2 \). Let \( j \) be the first integer for which \( q^{k_2} \neq 1 \mod p^j \). Such a \( j > 2 \) exists since \( q^{k_2} \neq 1 \) as integers so taking \( j \) such that \( p^j > q^{k_2} \) shows existence. We have \( k_2 = k_{j-1} \). We claim that \( pk_2 \) is the smallest integer \( l \) for which \( q^l = 1 \mod p^j \) and in addition \( q^{pk_2} \neq 1 \mod p^{j+1} \). We know by the definition of \( j \) that \( q^{k_2} = 1 \mod p^{j-1} \), i.e., \( q^{k_2} = 1 + rp^{j-1} \) for some integer \( r \) which is prime to \( p \). Taking the \( p^j \)th power we have

\[
q^{pk_2} = (1 + rp^{j-1})^p = 1 + rpp^{j-1} + dp^{2j-2} = 1 + rp^j + dp^{2j-2}
\]

where the term \( dp^{2j-2} \) corresponds to all the additional terms in the binomial expansion all of whom are divisible by \( p^{2j-2} \). Since \( j > 2 \), we have \( 2j-2 > j \) and so \( dp^{2j-2} = cp^{j+1} \) for some integer \( e \) and we see that \( q^{pk_2} \) satisfies the conditions of our claim. It remains to be seen that no integer \( l < pk_2 \) satisfies \( q^l = 1 \mod p^j \). Let us write \( l = fk_2 + s \) for some \( f < p \) and some \( r < k_2 \). We have \( q^l = (q^{k_2})^f q^s \). If \( s \neq 0 \) this expression will not equal 1 modulo \( p^2 \), let alone modulo \( p^j \). If \( r = 0 \) we have \( q^l = (1 + rp^{j-1})^f = 1 + rfp^{j-1} + gp^{j+1} \) which again does not equal 1 modulo \( p^j \) since neither \( r \) nor \( f \) is divisible by \( p \). We see that \( pk_j = k_{j+1} \) and that we can continue inductively using
the exact same argument to show that for $i > j$ we have $k_{i+1} = pk_i$, thus proving the lemma. \textit{q.e.d.}

We continue with the proof of the theorem. Consider the function

$$g_N(x) = \frac{1}{N} \sum_{k=0}^{N-1} e^{2\pi i q k x}$$

which appears in Weyl's equi-distribution criterion for the sequence $x(k) = q^k x$. We want to estimate $g_N(x)$ and show that it tends to zero for almost all $x$ w.r.t. $\mu$.

Given some $n$, let

\begin{equation}
\phi_n(x) = \prod_{i=0}^{n-1} \nu_i(x_i)
\end{equation}

i.e., $\phi_n(x)$ is the measure $\mu(I_j)$ for $j$ such that $x \in \left[\frac{j}{p^n}, \frac{j+1}{p^n}\right]$. Consider the measure

$$\omega_n(x) = \sum_{j=0}^{p^n-1} \mu(x + \frac{j}{p^n})$$

the sum of all level $n$ shifts of $\mu$. For any interval $I \subset I_k$ we have

$$\frac{\mu(I)}{\mu(I + \frac{j}{p^n})} = \frac{\mu(k)}{\mu(k + j \mod p^n)}$$

hence, we have

$$d\omega_n(x) = \frac{d\mu(x)}{\phi_n(x)}$$

Assume that in the notation of the lemma $N \leq k_n$. We get

$$\int \frac{|g_N(x)|^2}{\phi_n(x)} d\mu(x) = \int |g_N(x)|^2 d\omega_n(x) = \int \sum_{j=0}^{p^n-1} |g_N(x + \frac{j}{p^n})|^2 d\mu(x)$$

$$= \frac{1}{N^2} \int \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} e^{2\pi i (q^k-q^l)x} \left( \sum_{j=0}^{p^n-1} e^{2\pi i (q^k-q^l)j/p^n} \right) d\mu(x)$$

If $k \neq l$ we have $q^k \neq q^l \mod p^n$ and hence $\sum_{j=0}^{p^n-1} e^{2\pi i (q^k-q^l)j/p^n} = 0$ since the sum is invariant under multiplication by $e^{2\pi i (q^k-q^l)/p^n} \neq 1$. On the other hand, if $k = l$ we get

$$\sum_{j=0}^{p^n-1} e^{2\pi i (q^k-q^l)j/p^n} = \sum_{j=0}^{p^n-1} 1 = p^n$$

We conclude that

$$\int \frac{|g_N(x)|^2}{\phi_n(x)} d\mu(x) = \frac{p^n}{N}$$

By our assumption that all $\nu_i(j) < a$ and equation 19 we have for all $x$,

\begin{equation}
\phi_n(x)^{1/n} < a = e^{-H}
\end{equation}
Choose some integer \( s \) such that \( Hs > \log(p) \) and let \( t \) be some real such that \( 1 < t < \frac{Hs}{\log(p)} \). For each integer \( j \), let \( n_j \) be such that \( cp^{n_j-1} = k_{n_j-1} < j^s \leq k_{n_j} = cp^{n_j} \). With these choices we have

\[
\int \left| g_{j^s}(x) \right|^2 \phi_{n_j}(x) d\mu(x) = \frac{p^{n_j}}{j^s} \leq \frac{p}{c}
\]

Since \( t > 1 \) we obtain

\[
\sum_{j \geq 1} \int \left| g_{j^s}(x) \right|^2 \phi_{n_j}(x) d\mu(x) < \infty.
\]

We want to conclude from this that for almost all \( x \) w.r.t. \( \mu \) we have

\[
\lim_{j \to \infty} \frac{\left| g_{j^s}(x) \right|^2}{j^t \phi_{n_j}(x)} = 0
\]

Fix some \( \varepsilon > 0 \). For a given \( j \) let \( A_{j,\varepsilon} \) be the set of \( x \) for which \( \frac{\left| g_{j^s}(x) \right|^2}{j^t \phi_{n_j}(x)} \geq \varepsilon \). Obviously \( \int \frac{\left| g_{j^s}(x) \right|^2}{j^t \phi_{n_j}(x)} d\mu(x) \geq \varepsilon \mu(A_{j,\varepsilon}) \) hence

\[
\sum_{j \geq 1} \mu(A_{j,\varepsilon}) < \infty
\]

Let \( B_{\varepsilon} \) be the set of \( x \) such that \( \limsup \frac{\left| g_{j^s}(x) \right|^2}{j^t \phi_{n_j}(x)} \geq \varepsilon \). By the Borel-Cantelli lemma (lemma 23 of the appendix), we see that \( \mu(B_{\varepsilon}) = 0 \). Taking \( B = \bigcup_j B_{1/2} \), we see that \( \mu(B) = 0 \), but \( B \) is the complement of the set we are looking for.

Since \( \frac{t \log(p)}{s} < H \), we have

\[
\phi_{n_j}^{1/n_j} < e^{-t \log(p)/s} = p^{-t/s}
\]

We also have \( p^{n_j} = k_{n_j}/c \geq j^s/c \) or \( c \geq j^s p^{-n_j} \). Putting the two estimates together we obtain

\[
j^t \phi_{n_j} \leq (p^{-n_j})^{t/s} (j^s)^{t/s} \leq c^{t/s}
\]

From this and (21) we conclude that

\[
(22) \quad \lim_{j \to \infty} g_{j^s}(x) = 0
\]

for almost all \( x \) w.r.t. \( \mu \).

All that is left is to extrapolate the limit from the subsequence \( g_{j^s} \) to all \( g_n \). Let \( j^s \leq n < (j+1)^s \). We write

\[
g_n(x) = \frac{1}{n} \sum_{l=1}^{n} e^{2\pi ml}x
\]

\[
= \frac{1}{n} \sum_{l=1}^{j^s} e^{2\pi ml}x + \frac{1}{n} \sum_{l=j^s+1}^{n} e^{2\pi ml}x = \frac{j^s}{n} g_{j^s}(x) + \frac{1}{n} \sum_{l=j^s+1}^{n} e^{2\pi ml}x
\]
The first term goes to zero by (22). For the second term we exploit the fact that \( |e^{2\pi mq'x}| = 1 \), so

\[
\frac{1}{n} \sum_{l=j+1}^{n} e^{2\pi mq'x} \leq \frac{n - j^s}{n} \leq \frac{(j+1)^s - j^s}{j^s} = \left(\frac{j+1}{j} - 1\right)^s
\]

which goes to zero as \( n \) and hence \( j \) tend to infinity, as desired. By Weyl’s criterion the sequence \( q'x \) is equi-distributed and hence, \( x \) is \( q \)-normal. q.e.d.

Following the theorem, consider a situation in which we assume that the user will generate I/O using some, say binary (\( p = 2 \)), self similar I-bias model, where the probability of 1 is bigger than the probability of 0. Consequently, we populate the cache with addresses which have more 1s than 0s. If the cache has size \( n^{1-\delta} \) for some \( \delta > 0 \), none of these addresses (in the limit where \( h \to \infty \)) will be 2-normal. Suppose that in the end, the user generates I/O using a ternary (\( q = 3 \)) model, then Host’s theorem tells us that regardless of \( \delta \), the hit ratio will approach 0. This is in contrast to the situation in which the user uses another binary bias model that prefers 1s to 0s consistently. In that case for \( \delta \) small enough, the hit ration will approach 1, a sign of ”compatibility”, or ”correlation”. We also note that if the user uses a binary model which prefers 0s, then even if we have a cache of size \( cn \), \( 0 < c < 1 \), the hit ratio will approach 0. This is a form of ”negative correlation”. We conjecture, that if a cache of size \( cn \) is prepared for an input from a self-similar (I-model) \( p \)-ary bias model and the user uses a \( q - ary \) bias model, the hit ratio will approach \( c \), corresponding to ”no correlation”.

The problem of having ”independent” or ”negatively correlated” distributions explains why optimal algorithms, tailored for a specific type of I/O stream, are seldom used. They are simply not robust. Instead, one uses robust algorithms such as FIFO or LRU that tend to adapt to the input stream and usually provide reasonably good performance.

Notes for chapter 2

(1) A notion of entropy for measure preserving transformations was introduced by Kolmogorov and Sinai, [115, 164]. Consider a measure space \( (X, \Sigma, \mu) \) and a measure preserving transformation \( T \). Let \( A = (A_1, ..., A_k) \) be a partition of \( X \) into disjoint measurable sets. For a given integer \( n > 0 \), we consider the partition \( A_n \) consisting of the sets \( A_{i_1, ..., i_n} = \bigcap_{j=1}^{n} T^{-j}(A_{i_j}) \). The measure \( \mu \) induces a finite probability measure \( P_{A,n} \) on the index sets by the formula \( p(i_1, ..., i_n) = \mu(\bigcap_{j=1}^{n} T^{-j}(A_{i_j})) \). It can be shown that

\[
H(T, A) = \lim_{n \to \infty} \frac{1}{n} H(P_{A,n})
\]

exists. This result is known as the Shannon-McMillan theorem.
2. A Fractal Model for Storage System Activity

Definition 24. We define the Kolmogorov-Sinai entropy of $T$, denoted by $H(T)$, to be the supremum of the limits in (23) over all possible partitions $A$.

Definition 25. We say that a partition $A = (A_1, ..., A_k)$ is a generating partition of the dynamical system $(X, \Sigma, \mu, T)$ if $\Sigma$ is the smallest $\sigma$-algebra containing the sets $T^n(A_j)$ for all $n \in \mathbb{Z}$ and $1 \leq j \leq k$.

A theorem of Kolmogorov and Sinai, [164], states that if $A$ is a generating partition then $H(T, A) = H(T)$.

Let $\mu$ be a self similar $p$-ary bias model. Let $T$ be the multiplication by $p$ operation. Consider the partition $A_j = (\frac{i}{p^j}, \frac{i+1}{p^j}]$, $j = 0, ..., p-1$. The set $\cap_{j=1}^n T^{-j}(A_{i_j})$ is simply the interval $(\frac{i}{p^n}, \frac{i+1}{p^n}]$, where $i$ is given by the sequence $i_1, ..., i_n$. It is also clear that $A$ generates the Borel $\sigma$-algebra, so $H(T, A) = H(T)$. However, $H(T, A)$, is non other than $H(\nu)$. In fact the sequence $\frac{1}{n} H(P_{A,n})$ is constant in this case with that value. In this particular case, the arguments that we used in analyzing the hit ratio of the optimal static algorithm show that for $\mu$ almost all $x$, $\lim_{n} \phi_n(x)^{1/n} = e^{-H(T)}$, explaining our use of the letter $H$ in inequality (20). The discussion above also explains why we can consider the inequality as a "positive entropy" condition for measures that are not necessarily self-similar (invariant).

One of the motivations for Host’s paper was to reprove the following result of Rudolph, [153], which provides a partial answer to a conjecture of Furstenberg, [76], about the scarcity (rigidity) of measures which are invariant with respect to several "independent" commuting transformations.

Theorem 13. Let $p, q > 1$ be integers which are prime to each other. and let $\mu$ be a probability measure on the unit interval (circle). assume that $\mu$ is invariant w.r.t. both multiplication by $p$ and multiplication by $q$, ergodic w.r.t. the semigroup generated by both and that multiplication by $p$ has positive entropy w.r.t. $\mu$. Then, $\mu$ is Lebesgue measure.

Host’s theorem and its technique of proof may be viewed as an important step in the analysis of rigidity, a line of research which eventually led to the proof by E. Lindenstrauss of an important result, quantum unique ergodicity for Maass forms, see [121] for statements of the results and a historical survey.

(2) When discussing average seek distances, we computed the average distance between points in a self-similar (fractal) measure. The average distance in other fractal structures has been studied in the context of the famous towers of Hanoi problem. In this problem, there are $n$ disks (not necessarily disk drives) of different
sizes stacked on three vertical pegs. At any stage, the disks on each peg must be sorted according to size with the largest disk at the bottom of the peg. An allowable step is to move the top disk from one peg to another, but only if the resulting configuration is legal. In the classical case which is a favorite exercise in computer science intended to explain the concept of recursion, one is asked to move from a configuration where all the disks are on one peg to a configuration where all the disks are on another peg. It is easy to show that this will take $2^n - 1$ steps. Denote by $S$ the random variable which measures the minimal number of steps which are required to move from one uniformly random legal configuration to another. In [48, 95, 151] it is shown that $E(S) \sim 466 \frac{2^n}{885}$. The result is proved by mapping the problem to the problem of computing the average distance between points in the Sierpinski gasket. The Sierpinski gasket is a fractal which is obtained by initially taking an equilateral triangle with edge length 1. We remove from the triangle the interior of the triangle whose vertices are the mid-points of the edges of the original triangle. After the removal, we are left with three equilateral triangles on the top, left and right sides of the original. We repeat the process for each one independently. The Hanoi tower with $n$ disks is related to a discrete approximation of the Sierpinski gasket, where a configuration is mapped to a point in a discrete Sierpinski gasket by calling the pegs, top, left, bottom and specifying a configuration by the sequence describing the peg location of the largest disk, second largest disk and so on. In [151] the computation of the average distance is performed by considering a combinatorial structure known as a finite probabilistic automaton which keeps track of the recursions involved. The case of average seek time that we have described can be described in a similar way using a much simpler automaton with just two states.
CHAPTER 3

Disk scheduling, airplane boarding and Lorentzian geometry

1. The disk scheduling problem

Modern disk drives have the ability to queue incoming read and write requests and to service them in an out of order fashion.

In the disk scheduling problem we are given a batch of \( n \) queued requests and we wish to service them in an order which minimizes the total service time, or equivalently, in an order which minimizes the number of disk rotations required to service all \( n \) requests.

There are two basic questions which we can ask:

(1) Given a set of requests, what is the best order to serve them. Find an algorithm that will quickly compute the best order of service.

(2) Given a set of requests and the optimal ordering, how much time will it take the disk to service all the requests.

The disk scheduling problem with various assumptions has a long history of both experimental and theoretical work. At first, due to technological constraints, only algorithms that take seeks into account, while ignoring disk rotation were considered. For analysis of such algorithms see, [13, 51, 46, 52, 53, 49, 40, 60, 75, 82, 96, 141, 172, 188, 191] and for experimental results see, [187, 190, 78]. Later, when technology changed, algorithms that consider both radial seeks and the rotational position were introduced and studied, see [9] for theory and [159, 99, 173, 174] for experimental results. It turns out that in general the problems above are difficult to answer, so we concentrate on a simplified case which can perhaps provide us with some intuition.

We will assume that the seek distance function \( f \) is linear \( f(t) = ct \). In fact, to simplify the notation, we will assume that \( c = 1 \). The case of a general value of \( c \) can be easily deduced from this particular case. Thus, we set \( f(t) = t \). physically, this choice amounts to ignoring the effects of acceleration and deceleration of radial disk head motion. We assume that the disk head reaches maximal velocity instantaneously and similarly, can stop instantaneously. Obviously, both assumptions are unrealistic, but it is easier to tackle this toy model and we can hope that it will provide us with some insights.

In the linear case, Andrews, Bender and Zhang (ABZ) have found a simple algorithm for computing the optimal tour, [9]. We describe their
algorithm, but to better understand it we briefly discuss the notion of a partially ordered set first.

2. Partially ordered sets

**Definition 26.** A set $X$ with a relation $R \subseteq X \times X$ is said to be a partially ordered set, or a poset, if the following properties hold:

1. For all $x \in X$, $(x,x) \in R$.
2. For all $x,y \in X$, $(x,y) \in R$ and $(y,x) \in R$ implies $x = y$.
3. For all $x,y,z \in X$, $(x,y) \in R$ and $(y,z) \in R$, implies $(x,z) \in R$.

We also have the following definitions for some basic notions related to posets.

**Definition 27.** Two elements $x,y \in X$, are called comparable if either $(x,y) \in R$ or $(y,x) \in R$. Two elements $x,y \in X$ are called incomparable if they are not comparable. A poset $X$ is said to be linearly ordered if any two elements of $X$ are comparable. A chain $C$ in $X$ is a linearly ordered subset of $X$. An anti-chain $A$ is a subset of elements, such that any two distinct elements $x,y \in A$ are incomparable. We say that an element $x \in X$ is minimal if there is no $y \neq x$ such that $y \leq_R x$.

Given a finite partially ordered set $(X, \leq_R)$, there is a simple and well known procedure which simultaneously finds a decomposition of $X$ into a minimal number of anti-chains and a chain in $X$ of maximal size.

**The wave propagation process:**
Consider the set $X_1$ of minimal elements in $X$. Let $X_i$ be the set of minimal elements in $X - \bigcup_{j=1}^{i-1} X_j$. For each element of $x \in X_i$ construct pointers (arrows) to all elements $x' \in X_{i-1}$ for which $x' \leq_R x$. Stop when reaching a $k$ for which $X = \bigcup_{j=1}^{k} X_j$. We will say that $X_i$ is the wave front at step $i$ (time $i$).

We observe that the sets $X_i$ are all anti-chains since the set of minimal elements of any poset is always an anti-chain by definition of minimality. It is also easy to see that for each $x \in X_i$, $i > 1$ there must be an element $x' \in X_{i-1}$ such that $x' \leq_R x$, since otherwise $x$ would have been minimal w.r.t. $X - \bigcup_{j=2}^{i-1} X_j$ and would have belonged already to $X_k$, $k \leq i - 1$. This means that there are pointers (arrows) from any element in $X_i$, $i > 1$.

Let $X_k$ be the last nonempty set. We can construct chains of size $k$ by following the pointers from elements of $X_k$ all the way back to $X_1$. The maximal size of a chain is $k$ since any longer chain will have at least two elements in common with one of the anti-chains $X_i$, but by definition a chain and an anti-chain can only intersect in a single point, leading to a contradiction. Similarly any decomposition into anti-chains will have at
least $k$ anti-chains to cover the elements of a chain of size $k$. Therefore the $X_i$ form a minimal decomposition and a chain constructed via the pointers will be maximal.

We note that the process can sometimes be applied to infinite posets. In particular if all chains in a poset have length at most $k$, then the process will end after at most $k$ steps. If any chain which ends at a given element is finite then the sets $X_i$ can be indexed by ordinals and the origin of the wave propagation process may be traced back to G. Cantor’s work on ordinal arithmetic and well founded sets, however we will not require such sweeping generality.

2.1. Partially ordered sets and the disk scheduling problem.

We would like to restate the disk scheduling problem in terms of partially ordered sets. For that, we need to find some meaningful partial orders in the context of disk drives. Consider the infinite band $U$ given by, $-\infty < t < \infty$, $0 \leq r \leq 1$. We think of the $t$ coordinate as representing time, in units of complete disk revolutions.

We will denote by $C$ the set of disk locations, in polar like coordinates $(\theta, r)$, $0 \leq \theta, r \leq 1$, that were introduced in chapter 1. Recall that we identify $(0, r)$ with $(1, r)$ and that we normalize the time coordinate in such a way that the head of the disk is at angle 0 at time 0. It follows that $\theta = t \pmod{1}$ is the angle of the disk head at time $t$.

We have a map $\pi : U \to C$ given by

$$\pi(t, r) = (t \pmod{1}, r) = (\theta, r)$$

If the disk head at time $t$ is in radial position $r$ then the disk position of the head is $\pi(t, r)$.

We have a mapping $T : U \to U$ which preserves disk locations, namely, $T(t, r) = (t + 1, r)$. We can think of $T(t, r)$ as representing the same location, one disk revolution later.

**Definition 28.** We define the horizontal partial order $\leq_{\text{hor}}$ on $U$. We say that $(t_1, r_1) \leq_{\text{hor}} (t_2, r_2)$ if and only if $t_2 \geq t_1$ and $t_2 - t_1 \geq |r_2 - r_1|$. If $f(t) = t$ is the seek function of the disk we may interpret this relation as follows, $(t_1, r_1) \leq_{\text{hor}} (t_2, r_2)$ if and only if the disk head which at time $t_1$ is in disk location $\pi(t_1, r_1)$ can reach at time $t_2$ the disk location $\pi(t_2, r_2)$.

Consider a set of I/O requests to data locations, $\bar{R} = \{R_1, ..., R_n\}$ on the disk. If we follow the radial position of the head of the disk over time we will trace out a curve in $U$ which by definition of the seek function and the horizontal partial order, will be a continuous chain in the horizontal partial order. It is also true by the definitions, that each curve which forms a chain w.r.t. $\leq_{\text{hor}}$ describes a possible motion of the head of the disk in our disk drive model. We thus equate such curves with possible disk head motions.

Each requested location $(\theta, r)$ can be visited at a time $t$ such that $t \pmod{1} = \theta$, or $(t, r) \in \pi^{-1}(\theta, r)$.
Definition 29. A tour of $\bar{R}$ is a chain in the partial order $\leq_{\text{hor}}$ starting at some point $(0, r_0)$ and containing one element from each set $\pi^{-1}(R_i)$, for all $i = 1, \ldots, n$. A tour describes the motion of the head of the disk as it services all the I/O requests, starting at time $t = 0$.

The service time of the set of requests $\bar{R}$, denoted $ST(\bar{R})$ is the minimal value of $t_{\text{finish}}$ for which there is a tour of $\bar{R}$ ending at some point of the form $(t_{\text{finish}}, r_{\text{finish}})$. The service time is the minimal number of disk rotations needed to service the requests in $\bar{R}$.

A tour with a minimal time span $t_{\text{finish}}$ among all tours is called an optimal tour.

Our restated goal is to find a tour which is either optimal or very close to optimal. We also want to provide a good estimate for the finishing time of such a tour.

3. The ABZ algorithm

For the purpose of finding an optimal tour, we consider a second partial order on $U$.

Definition 30. We let $(t_1, r_1) \leq_{\text{ver}} (t_2, r_2)$ if and only if $|t_2 - t_1| \leq |r_2 - r_1|$ and $r_2 \geq r_1$. We call this partial order the vertical partial order on $U$.

For simplicity, we shall always assume that for any pair of requests $(t_1, r_1), (t_2, r_2)$ we have $|t_1 - t_2| \neq |r_1 - r_2|$. Under this assumption any chain in $\leq_{\text{ver}}$ is an anti-chain in $\leq_{\text{hor}}$ and conversely, any chain in $\leq_{\text{hor}}$ is an anti-chain in $\leq_{\text{ver}}$. In this sense the two partial orders are complementary.

We can now describe the algorithm of Andrews, Bender and Zhang, [9], which finds an essentially optimal tour.

Theorem 14. Let $\bar{R}$ be a set of requested data locations and $\bar{R}^U = \pi^{-1}(\bar{R})$ the set of points in $U$ which represents these data locations. Assume that the size of a minimal decomposition of $\bar{R}^U, \leq_{\text{ver}}$ into anti-chains is of size $m$. Then, we can effectively find a tour $\tau$ such that $t_{\text{finish}} \leq m + 1$. In addition, any tour satisfies $t_{\text{finish}} \geq m - 2$, hence $\tau$ is at most three disk rotations from being optimal regardless of the number of requests.

Proof: Let $W_1, \ldots, W_m$ be the sets obtained in the wave propagation process applied to the partially ordered set $(\bar{R}^U, \leq_{\text{ver}})$. The sets $W_i$ are anti-chains with respect to $\leq_{\text{vert}}$ and so by complementarity, they form chains w.r.t. $\leq_{\text{hor}}$. Consequently, the head of the disk can move along $W_i$. The chain structure provides a linear order on $W_i$. By the properties of the wave propagation process the number of anti-chains equals the maximal size of a chain in $\bar{R}^U$ w.r.t. $\leq_{\text{ver}}$. The translation map $T$ is an $\leq_{\text{vert}}, \leq_{\text{hor}}$ preserving automorphism of $\bar{R}^U$, hence, the sets $W_i$ are periodic with period 1, i.e., $T(W_i) = W_i$, and hence $W_i$ have the form $\pi^{-1}(S_i)$ for some $S_i \subset \bar{R}$. 


Consider the piecewise linear curves $L_i$, $i = 1, \ldots, m$, in $U$ which join successive points of $W_i$ by straight line segments. Since the $W_i$ are chains, and the seek function is linear, the curves $L_i$ form continuous chains with respect to $\leq_{\text{hor}}$. The piecewise linear curves $L_i$ can be considered as graphs of functions on $t$ and by abuse of notation we will denote by $L_i(t)$ the value of the function at $t$. If we think of $L_i(t)$ as describing the motion of the head of the disk, then $L_i(t)$ is simply the radial location of the head at time $t$. We note that since $T(W_i) = W_i$, we also have $T(L_i) = L_i$.

Let $J$ be the line segment in $U$ given by the equation $r = t$, for $0 \leq t \leq 1$. Since $J(0) = 0 \leq L_i(0)$ and $J(1) = 1 \geq L_i(1)$, there exist times $t_i$ where the piecewise linear curves $L_i$, meet the segment $J$. For simplicity we assume that $t_1 \leq t_2 \leq \ldots \leq t_m$, otherwise we can re-order the curves $L_i$ so that this property holds. We let $u_i = (t_i, L_i(t_i))$ be a meeting point of $J$ and $L_i$. We construct a tour as follows.

**The ABZ algorithm:** The tour starts at $u_0 = (0,0)$. The tour follows the curve $J$ until the point $u_1$. It then follows the curve $L_1$ for an entire disk rotation, ending up at the point $T(u_1)$. We continue by following the curve $T(J)$, which is $J$ shifted by 1 to the right until we encounter the point $T(u_2)$ which is also on the curve $L_2 = T(L_2)$. We follow $L_2$ for an entire disk rotation until the point $T^{2}(u_2)$, which is the point $u_2$, shifted two units to the right. The tour proceeds in the same manner from $T^i(u_{i+1})$ along $L_{i+1}$ for 1 time unit (a single disk rotation), getting to $T^{i+1}(u_{i+1})$ and then moving along $T^{i+1}(J)$ to $T^{i+1}(u_{i+2})$.

The tour ends at $T^{m}(u_m)$. Since $u_m$ has time coordinate $t$ which satisfies $0 \leq t \leq 1$ and $T^{m}$ shifts time by $m$ units we conclude that the total time for the tour is at most $m + 1$. It does pass through all requested locations since it passes through the curve $L_i$ for an entire revolution, passing through each disk location of the curve $L_i$, which is periodic with period 1.

To complete the proof, we claim that any tour requires at least $m - 2$ disk rotations. To see the lower bound, let $Q_1 = (t_1, r_1)$ and $Q_2 = (t_2, r_2)$ be any pair of points in $U$ such that $Q_2 \geq_{\text{ver}} Q_1$. Let $Q_2' = (t_2', r_2') \in U$ be any point which represents the same disk location as $Q_2$, namely, $\pi(Q_2') = \pi(Q_2)$ and such that $Q_2' \geq_{\text{hor}} Q_1$.

If $t_2 \geq t_1$ then by definition of the relation $\leq_{\text{ver}}$ we have $(t, r_2) \geq_{\text{ver}} Q_1$ for all $t_1 \leq t \leq t_2$. We conclude by complementarity that non of these points satisfies $(t, r_2) \geq_{\text{hor}} Q_1$ and therefore $t_2' \geq t_2 + 1 = t_1 + (t_2 - t_1) + 1$. If $t_2 \leq t_1$ then obviously $t_2' \geq t_2 + 1$ since our assumption requires that $t_2' > t_1$. The same analysis works to show that $t_2' \geq t_2 + 1$ when we assume the reverse relation $Q_1 \geq_{\text{ver}} Q_2$.

Let $Q_1, \ldots, Q_m$ be a maximal chain in $R^U$ w.r.t. $\leq_{\text{ver}}$. Consider an optimal tour in $U$. Assume that the points $Q_i$ are indexed by the order in which the optimal tour visits their equivalence classes w.r.t. $T$. This does not coincide necessarily with their order in the chain. Let $Q_i'$ be the point
56 3. DISK SCHEDULING, AIRPLANE BOARDING AND LORENTZIAN GEOMETRY

in the equivalence class of \( Q_i \) which the tour visits. Consider \( Q_1 \) and \( Q_2 \). By the chain condition we have either \( Q_1 \geq \text{ver} \ Q_2 \) or vice versa. In any case we conclude from the previous argument that \( t'_2 \geq t_1 + (t_2 - t_1) + 1 \). More generally, assume that \( Q'_i = T^i_1(Q_i) \). Applying the argument above to \( Q'_i \) and \( T^i_1(Q_{i+1}) \) we conclude that \( t'_{i+1} \geq t'_i + (t_{i+1} - t_i) + 1 \), which leads to

\[
t'_m - t'_1 \geq (m - 1) + \sum_{i=1}^{m-1} t_{i+1} - t_i = (m - 1) + (t_m - t_1)
\]

Since \( Q_1 \) and \( Q_m \) are comparable w.r.t. \( \leq \text{ver} \) we have \( t_m - t_1 \geq -1 \) which leads to the desired lower bound. \( q.e.d. \)

We seem to have solved our basic problem, but we should explore a little further. One issue is, how quickly can we find the ABZ tour, or in other words, compute the wave fronts \( W_i \)?

### 3.1. Efficiently computing the ABZ algorithm

We first recall that \( W_i \) are periodic in the \( t \) direction with period 1, therefore it is enough to compute \( W_i \) restricted to the set of points \( S \) with \( 0 \leq t \leq 1 \). Consider a point \( Q \in \bar{R}^U \). The set \( W_i \) to which \( Q \) belongs is determined by examining the set of elements \( I^- (Q) = \{ R \in \bar{R}^U \mid R \leq \text{ver} \ Q \} \), in fact this is generally true for the wave propagation process on any partially ordered set. If \( Q \in S \) then any \((t,r) \in I^- (Q)\) satisfies \(-1 \leq t \leq 2 \). We let \( \tilde{S} \subset U \) be the set of points satisfying this property. We can therefore restrict the wave propagation process to \( X = \bar{R}^U \cap \tilde{S} \), a set which contains at most \( 3n \) points.

To compute the wave propagation process in a general finite partially ordered set, we can make a list for each element \( Q \), of the elements in \( I^- Q \), which are not \( Q \) itself. The elements in \( W_1 \) will be those with empty lists. Then we take each element of \( W_1 \) and remove it from all lists and repeat the process. In more detail. Assume that the poset has \( m \) elements and enumerate the elements of the poset \( 1, \ldots, m \). Given element \( i \) compare it to all other elements according to their enumeration and place them in \( I^- (i) \) if needed. In the process we count the number of elements in \( I^- (i) \). All this takes roughly \( c_1 m^2 \) steps for some constant \( c_1 > 0 \), since we need to compare all pairs. Checking for empty sets to determine the next wave front can take up to \( c_2 m \) steps since we know how many elements there are in each set \( I^- (i) \). When we remove an element, we have to decrease the number of elements in \( I^- (i) \) for each set it belongs to. We can find using binary search, whether it belongs to \( I^- (i) \) in at most \( \log(m) \) steps. Repeating for each element removed and each \( i \) can lead to at most \( c_3 m^2 \log(m) \) steps and in some cases we will actually have to perform that many steps, for example, in the case of a linearly ordered set. All in all the process may take up to \( c_4 m^2 \log(m) \) steps.

#### 3.1.1. Patience sorting

For the particular posets which are of interest to us the wave propagation process can be computed much more effectively.
3. THE ABZ ALGORITHM

The first step is to convert the partial order \( \preceq \) to a new partial order \( \preceq_{pos} \).

**Definition 3.1.** If \( v_1, v_2 \in \mathbb{R}^n \) are vectors, let \( v_1 \preceq_{pos} v_2 \) if and only if all the coordinates of the vector \( v_2 - v_1 \) are non-negative. We will call it the **Positive partial order**. If \( n = 2 \) this is the same as saying that the vector \( v_2 - v_1 \) is in the positive quadrant.

In the \( n = 2 \) case it is easy to see that a 45 degree clockwise rotation converts \( \preceq \) into \( \preceq_{pos} \). Our assumption that no pair of points in \( \mathbb{R}^U \) satisfies \( |t_1 - t_2| = |r_1 - r_2| \) translates into the condition that no pair of images (after rotation) have the same standard coordinates which we will call the \( x \) and \( y \) coordinates instead of \( t \) and \( r \).

We are thus reduced to computing the wave propagation process for \( \preceq_{pos} \). For a finite set \( Y \) in the plane with partial order \( \preceq_{pos} \), satisfying the above condition, we can compute the wave propagation process more efficiently using a process known as Patience sorting which is described as follows.

**Patience Sorting algorithm:** We sort the elements according to their \( x \) coordinate and think of them as cards in a deck. The point with the smallest \( x \) coordinate corresponding to the top card in the deck, the point with the second smallest \( x \) coordinate corresponding to the second card from the top and so on.

On each card of the deck we write the \( y \) coordinate of the point corresponding to it. Next, we will need a big table. We lay down the cards from the deck into piles on the table according to the following rule.

The first card is laid out at the left end of the table. At any given time, the cards that have been laid out are in several piles on the table. The new card to be laid out goes to the leftmost pile such that the value of its top card is larger than the value of the laid out card.

We claim that at the end of this process, the leftmost pile contains the cards whose corresponding points are in \( W_1 \), the wave front at time 1 w.r.t. \( \preceq_{pos} \), the second pile from the left, those corresponding to points in \( W_2 \), the wave front at time 2, and so on. To see this, it is enough to convince ourselves that this is true of the leftmost pile, since the process restricted to the other piles is the same as the one obtained by removing the cards of the first pile from the deck. Assume that card \( A \) has ended up in the leftmost pile. If the corresponding point is not in \( W_1 \), it means that there is another point with smaller \( x \) and \( y \) coordinates. A smaller \( x \), means that the corresponding card \( B \) was already laid out on the table, say at the \( k' \)th pile from the left. Since the value of the topmost card of a pile only decreases with time, the value of the top card of the \( k' \)th pile at the time of arrival of card \( A \) was smaller or equal to the value of card \( B \). It is easy to check inductively that at each stage the values of the topmost cards are sorted in
increasing order from left to right. Since the value of $B$ is smaller than that of $A$, this means that card $A$ could not have ended up in the first $k$ piles from the left contrary to our assumption. Also, a card corresponding to a point in $W_1$ ends up in the first pile since its value is smaller than those of all preceding cards and we are done.

Patience sorting was invented in 1963 by Mallows, [127],[128] as a practical method to humanly sort an actual deck of cards. The process is algorithmically fast. For each card we need to determine its pile. Since at each stage the values of the topmost cards are sorted from left to right, a binary search will find the correct pile. Consequently it takes an order of magnitude of $m \log(m)$ steps to complete the process, a good improvement over the previous method. For much more information on patience sorting we refer the reader to the excellent survey paper, [2].

3.2. Parallel computation via airplane boarding. There is another way to compute the wave propagation process on $\leq_{pos}$, this time not with cards, but with actual people. You will also need an airplane. Each point will be assigned a person. Again we sort the points according to their $x$ coordinate from smallest to largest. We place people in a queue to board the airplane, with the first person corresponding to the point with smallest $x$ coordinate, the second in line to the point with the second smallest $x$ coordinate and so on. We now assign each passenger, a row in the airplane. The row is given by the rank of the $y$ coordinate of the corresponding point, namely, if the $y$ coordinate is the $k$ smallest among all $y$ coordinates, the passenger is assigned to row $k$.

We make a few assumptions on the nature of the boarding process. We assume that passengers can walk very quickly one after the other along the narrow aisle of the airplane as long as someone is not blocking them. We also assume that passengers are infinitely (cardboard) thin. This disturbing assumption will be discussed later. There is a happier alternative. We can assume that the airplane is very spacious with ample (almost infinite) leg room between the rows.

When a passenger gets to their assigned row, it takes them a fixed amount of time to get organized. During that time they are standing in the aisle and blocking other passengers from getting through. After getting organized, they sit down, clearing the aisle for other passengers passing by on their way to their seat.

The assumption that passengers are infinitely thin means that many passengers can line up behind a blocking passenger without blocking any of the passengers which need to get to earlier rows. We understand that this is unrealistic, but we will deal with that later on.

If we consider the boarding process as it proceeds according to this model we see again the wave fronts being computed in parallel by the passengers. The set $W_1$ consists of of all passengers which proceed to their designated
rows immediately, since no passenger is blocking them from getting there. The set \( W_2 \) consists of all passengers which can proceed immediately to their row once the passengers from the set \( W_1 \) have taken their seats, an event which takes place simultaneously according to the model, since the delay was fixed.

We see that the passengers get seated in rounds (steps) with the passengers in \( W_i \) getting seated in round/step \( i \).

We can say that the speed of the calculation, which is the boarding time is equal to the number of rounds. This is obviously at most \( m \). We gained speed over the previous process by exploiting parallelism, there are many passengers and they can seat themselves simultaneously in some cases.

4. The general airplane boarding process

The description of airplane boarding with infinitely thin passengers suggests that airplane boarding which is an interesting performance analysis problem in its own right (actually more interesting than disk scheduling), may be related to disk scheduling and perhaps the two can be studied together. Suppose we want to consider airplane boarding more generally, without the unrealistic assumption that passengers are infinitely thin. In addition, our boarding model assumed a single passenger per row, while in reality there are more. We will attempt to model the airplane boarding process without these restrictions.

The input data for the model is composed of the following items.

- A number \( n \) which represents the number of passengers.
- A number \( h \) which represents the number of passengers per row.
- A sequence of passengers, indexed, 1, \ldots, \( n \), according to their position in the boarding queue, from first to last. Passenger \( i \) has a seat in an assigned row, denoted by \( r_i \). The row number \( r_i \) ranges between 1 and \( m = n/h \), and each number in that range is the row number of \( h \) different passengers. We implicitly assume here that the airplane is full.
- A length value \( w \), which measures the aisle length occupied by a passenger, baggage and personal space included.
- A length parameter \( l \), representing the distance between successive rows.

In addition to the above we will assume a fixed delay value \( d \) which we normalize by setting \( d = 1 \). The delay value measures the total amount of time that passes from the moment the passenger has reached his/her designated row until they clear the aisle. This time includes getting organized, placing carry-on luggage and, possibly, passing by previously seated passengers from the same row on the way to the designated seat. The last operation usually requires the seated passengers to get up and sit back after the newly arrived passenger has taken his/her seat.
We remark that the model could have been more detailed. For example, both the width \( w \) and the delay \( d \) could have been passenger dependent. The length parameter \( l \) could vary by passenger class, first, business or traveler and so on. Such models have also been considered and analyzed, however, for the sake of simplicity we shall restrict ourselves to the scenario above. As it turns out, the existing model captures the essence of the process.

Given the input data, the boarding process proceeds as follows. The location of row \( j \) along the aisle is \( j_l \). In each step, passengers advance as far as they can along the aisle towards their assigned row. When passenger \( i \) reaches his/her row, \( r_i \), he/she occupies the aisle space segment from \( l r_i + W \). All passengers that have reached their row take one time unit to get organized and sit down. Once they clear the aisle, passengers immediately behind them advance along the aisle as far as they can, towards their row and the process repeats. The boarding time \( T \) is the number of steps needed for all passengers to sit down.

In terms of these parameters, we define a precedence/blocking partial order relation. We begin by recalling a general procedure on relations

**Definition 32.** Given a relation \( R \) on a set \( X \), the **transitive closure** of \( R \) is the relation \( R^c \) such that \( (x, y) \in R^c \) if and only if there exist \( x = x_1, \ldots, x_k = y \in X \) such that \( (x_i, x_{i+1}) \in R \) for all \( 1 \leq i \leq k - 1 \).

Let \( i_1, \ldots, i_g, j \) be passengers who are lined up behind passenger \( i \) (since they cannot reach their row) just before he/she sits down and clears the aisle.

**Definition 33.** We say that passenger \( i \) **blocks** passenger \( j \), if

\[
(24) \quad l r_i - g w < l r_j.
\]

The **airplane boarding partial order** \( \leq_{\text{boar}} \) is the transitive closure of the blocking relation.

The blocking relation can be explained as follows. The \( g \) passengers lined between passengers \( i \) create an aisle backlog which stretches back to row \( l r_i - g w \), and if the inequality holds, this prevents passenger \( j \) from getting to his/her row.

In terms of the boarding partial order, the boarding process is simply the wave propagation process, with the wave front \( W_i \) consisting of passengers who have sat down after \( i \) time steps.

As before we can attach an airplane boarding process to \( n \) points in the plane. We determine the boarding queue location by sorting the points according to their \( x \) coordinate, while the ordering according to the \( y \) coordinate determines the passenger row. The \( h \) points/passengers with the lowest value of the \( h \) coordinate are assigned to row 1, the next \( h \) to row 2 and so on. However, the boarding partial order will no longer coincide with \( \leq_{\text{pos}} \) on the points, it will be somewhat more complicated. In fact, the boarding relation between two points \( i, j \) depends on many other points and cannot be determined just by knowing the coordinates of \( i \) and \( j \).
Conversely, given \( n \) passengers, lined up in a boarding queue we can assign to passenger \( i \) the normalized queue location \( q_i = i/n \) and the normalized row location \( r_i/m \), thus we can represent passengers as points in the unit square.

5. Disk scheduling and airplane boarding in a probabilistic setting

Having gained some connection between disk scheduling and airplane boarding, new questions come to mind.

Consider the boarding process in an airline that does not employ a boarding policy, namely, passengers are free to join the boarding queue whenever they wish. If we look at the order in which passengers boarded the airplane with respect to their row locations we will see that each time the order is different.

A passenger boarding a specific airplane will be interested naturally in the specific boarding time of his/her flight, but from the point of view of the airline, the interest is in the statistical behavior of the boarding time for randomly ordered passengers. In addition, airlines employ boarding policies. For example, the airline may first board passengers from rows 40 and above, followed by passengers from rows 30 and above, not to mention first class passengers which usually board before all others.

Such policies affect the order in which passengers board the airplane. Again, the effect is statistical, the policy does not determine the ordering, it just affects the ordering. The airline might want to know how the policy, affects the statistics of boarding time. After all, the main motivating factor (apart from preferred customer treatment) behind the policies is to try to lower boarding time, so the airplane can fly on time.

We can model the airline policy by considering a joint distribution \( \mu_p = p(q,r)dqdr \) on the unit square, with points describing the normalized queue and row positions of passengers. We generate a boarding process by sampling \( \mu_p \), \( n \) times independently, producing \( n \) points, which are translated into passenger queue positions and rows as described in the previous section.

We note that the translation into row and queue positions depends only on the ordering of points according to the \( q \) and \( r \) coordinate. Consequently, a change of variable of the form \((q, r) \rightarrow (f(q), g(r))\), where \( f, g \) are increasing functions, with \( f(0) = g(0) = 0 \) and \( f(1) = g(1) = 1 \) does not change the end result. We consider distributions which are obtained one from the other using such transformations as being equivalent. Given some \( \mu_p \), using such coordinate changes we can obtain a unique distribution \( \mu_{\tilde{p}} \), equivalent to \( \mu_p \) such that for any \( 0 \leq q \leq 1 \) we have \( \int_0^1 \tilde{p}(q,r)dr = 1 \) and for any \( 0 \leq r \leq 1 \) we have \( \int_0^1 \tilde{p}(q,r)dq = 1 \). From now on we assume that \( \mu_p \) is normalized in this way.
To see how we model an airline policy using a joint queue/row location density we consider a popular family of airline boarding policies.

Many airlines try to board passengers from the back of the airplane towards the front. In a typical example, they may first restrict boarding to passengers who are sitting in a back section of the airplane, say from rows 40 and above. After these passengers line up, they call passengers from, say, row 30 and above, then passengers from row 15 and above and finally all passengers.

We will describe such policies by considering the portion of passengers in each boarding group.

**Definition 34.** A **back-to-front boarding policy** is given by a monotone decreasing sequence of numbers \( \bar{r} = (r_0, r_1, \ldots, r_m) \), \( 1 = r_0 > r_1 > \ldots > r_{m-1} > r_m = 0 \). The sequence \( \bar{r} = (r_1, \ldots, r_{m-1}) \) is referred to as a **partition of size** \( m \). Assume that the airplane has \( s \) rows. The set of passengers who are seated between rows \( [r_i - 1 \cdot s] \) and \( [r_i \cdot s] \) is called the \( i \)’th group of passengers. The back-to-front policy corresponding to the partition allows the passengers from the first group of rows to join the queue first, followed by passengers from the second group and so on. We will denote this policy by \( F_{\bar{r}} \). The back-to-front policy with \( m = 1 \) which has a single group encompassing all passengers is called the **random boarding policy** and will be denoted by \( F \).

We show how to assign to a back-to-front policy a joint density function \( p(q, r) \), which describes the probability that a passenger sitting in row \( r \) will have queue position \( q \). We assume that the airplane is full and that passengers abide by the boarding policy, namely, passengers from group 1 queue before those of group 2 and so on. We assume that the order within the group is arbitrary since it was not specified by the policy. Under these assumptions in a back-to-front policy with parameters \( \bar{r} \), passengers in the \( i \)th group occupy positions \( (1 - r_{i-1})n \) to \( (1 - r_i - 1)n \) in the queue, therefore, the normalized queue/row coordinates of passengers \( (q, r) \) in the \( i \)th group satisfy \( r_{i-1} \geq r \geq r_i \) and \( 1 - r_{i-1} \leq q \leq 1 - r_i \). We denote the square given by these inequalities by \( S_i \). The set of squares \( S_i, i = 1, \ldots, m \), contains the anti-diagonal segment given by \( q + r = 1, 0 \leq q, r \leq 1 \). Since a passenger in a row \( r \) is equally likely to have any of the allowable queue positions, the probability density function \( p \) is defined by \( p(q, r) = 1/(r_{i-1} - r_i) \) if \( (q, r) \in S_i, i = 1, \ldots, m \), and \( p(q, r) = 0 \) otherwise (outside the squares \( S_i \)). In the case of the random boarding policy we see that \( p(q, r) = 1 \).

Passenger queues which are generated by sampling points using this distribution will correspond well with the policy restrictions. In fact, if we would sample precisely \( (r_{i-1} - r_i)n \) points from each square \( S_i \) then all passengers from group 1 will line before all passengers from group 2 and so on. Due to fluctuations in the number of sampled points from each region \( S_i \), some passengers may miss their intended queue range, but that is fine, passengers do not strictly adhere to the airline policy anyway.
Figure 1. A graphic illustration of a partition into five boarding groups of different sizes. Each square corresponds to one group. The bottom edges $B_1, B_2, \ldots, B_5$ of the squares $S_1, S_2, \ldots, S_5$, respectively, are depicted by a solid thick line. This is a partition of size 5.

Following the model above we can define boarding time as a random variable for a given airline policy and airplane configuration as follows:

**Definition 35.** Let $\mu_p = p(q, r) dqdr$ be a joint queue/row location distribution representing an airline boarding policy. Consider an airplane with $m$ rows and $h$ seats per row. Let $n = mh$ denote the number of passengers in a full airplane. Let $l$ denote the distance between successive rows (leg room) and let $w$ denote the aisle length that a single passenger occupies while standing in the aisle (personal space). The **boarding time** $L_{n, p, h, l, w}$ is the random variable given by the longest chain in the boarding partial order of passengers whose queue and row positions are determined by $n$ points sampled i.i.d. from the unit square using the density $p$.

Similar considerations can be made in the context of disk scheduling. Some areas of the disk are more popular than others, hence there will be more requests in these areas, however it is unlikely that the same batch of requests will repeat itself. Therefore, instead of looking at the service time for a particular batch of requests, we should consider the statistical properties of service time, given a certain request distribution.

Towards that end, we introduce a density function $p(r, \theta)$ which models the popularity of data at location $(r, \theta)$. When we choose the requests in
a batch, we assume that they are sampled independently from the measure 
\[ \mu_p = p(r, \theta) dr d\theta. \]

**Definition 36.** Let \( \mu_p = p(r, \theta) dr d\theta \) be a measure of the popularity of disk locations. The **service time** random variable \( L_{n,p} \) is defined as the size of the longest chain w.r.t. \( \leq_{\text{ver}} \) in the set \( \bar{R}^U \), where \( \bar{R} \) consists of \( n \) data locations in the disk \( C \) sampled independently using the measure \( \mu_p \). According to theorem 14 this is essentially the service time of the requests with an optimal tour.

Our re-stated goal is to understand the statistical behavior of \( L_{n,p,h,l,w} \) and of \( L_{n,p} \). The number of passengers per aisle in an airplane may be about 200, while the number of batched disk I/O requests may be a few dozen. These values of \( n \) are not particularly large, but are not too small either. As with many problems in probability it is actually easier to understand the behavior of \( L_{n,p,h,l,w} \) as the number of passengers \( n \) goes to infinity. The same is true for \( L_{n,p} \) as the number of requests becomes large. Consequently, our strategy will be to consider the asymptotic behavior of boarding and service times as \( n \) becomes large, while all other parameters remain fixed. Hopefully we can gain some insights from these computations also for some smaller values of \( n \).

We have seen that the airplane boarding and disk scheduling problems are connected. We may ask whether there is a more general context which will allow us to analyze statistical properties of disk scheduling and airplane boarding simultaneously, a kind of two for the price of one theory that will cover both as special cases. As it turns out there is such a theory and its roots go back in time to an era before disks or commercial airplanes existed. This theory will give us a 3 for the price of one deal, it will enable us to model the statistical properties of disk scheduling, the statistics of airplane boarding and the universe. Oddly, it does not seem to model many other things. The objects in this theory will be discrete space-time. These serve as discrete approximations to continuous objects known as space-time manifolds which are the basic objects in Lorentzian (space-time) geometry, the mathematical theory that models relativity theory.

The next section will introduce some basic concepts and facts about Lorentzian geometry and its discrete analogue. The discussion will be very basic, without explicit discussion of the t-word (tensor) nor the c-words (curvature, connection). We have placed all proofs in the corresponding section of the appendix. We refer the reader to [142] for a much more comprehensive and sophisticated treatment of the basic (continuous) Lorentzian geometry that we introduce.

### 6. Lorentzian geometry

**6.1. Minkowski space.** We consider the diagonal matrix \( D_{n,1} \), whose first diagonal entry is 1 and whose last \( n \) diagonal entries are \(-1\). We recall the following definition
For \( M = D_{n,1} \), letting \( u = (t, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} \), the quadratic form associated with \( M \) is given by
\[
Q_M(u) = u D_{n,1} u^t = t^2 - x_1^2 - \ldots - x_n^2
\]
We can consider any invertible linear transformation \( T \) applied to the space \( \mathbb{R}^{n+1} \). The quadratic form associated with \( D_{n,1} \), will be transformed into the quadratic form associated with the matrix \( M = T D_{n,1} T^t \). The matrix \( A \) will be symmetric and will have one positive eigenvalue and \( n \) negative eigenvalues. By a classical theorem of Sylvester, any matrix \( M \) with these properties can be obtained in this way.

**Definition 37.** Let \( M \) be a matrix of the form \( M = T D_{n,1} T^t \), where \( T \) is invertible. We will call such a matrix a **Minkowski matrix**. The **Minkowski metric** associated with \( M \) is the function \( q_M : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \rightarrow \mathbb{R} \) given by
\[
q_M(v, w) = Q_M(v - w)
\]
The vector space \( \mathbb{R}^{n+1} \) equipped with a Minkowski metric will be called a **Minkowski space**. If \( M = D_{n,1} \) the associated Minkowski space will be called the **standard Minkowski space**. Points in Minkowski space will sometimes be called **events**.

By abuse of notation we will usually denote the Minkowski space whose Minkowski matrix is \( M \) also by the letter \( M \).

**6.1.1. Classification of vectors and the causal partial order.** The quadratic form associated with \( M \) defines several basic types of vectors.

**Definition 38.** We say that a vector \( v \in \mathbb{R}^{n+1} \) is **causal** if \( v M v^t \geq 0 \). We say that a vector \( v \) is **time-like** if \( v M v^t > 0 \). We say that \( v \) is **light-like** if \( v M v^t = 0 \). We say that \( v \) is **space-like** if \( v M v^t < 0 \).

Consider the classification for the standard case of \( M = D_{n,1} \). The set of causal vectors \( v = (t, x_1, \ldots, x_n) \) consists of vectors that satisfy
\[
t^2 - \sum_{i=1}^{n} x_i^2 \geq 0
\]
This set separates into two sets \( C^- \) and \( C^+ \) consisting of vectors with \( t \leq 0 \) and those with \( t \geq 0 \) which intersect only at the origin. As sets we have \( C^- = -C^+ \). Geometrically, these sets are cones consisting of all rays (half lines) that connect the origin with an \( n \) dimensional ball. In the case of \( C^+ \) the ball \( B^+ \) is the set of vectors of the form \( (1, x_1, \ldots, x_n) \) with \( \sum_{i=1}^{n} x_i^2 \leq 1 \). The picture is the same for a general Minkowski metric \( M = L D_{n,1} L^t \), with the two cones being the images of \( C^- \) and \( C^+ \) via the linear transformation \( L \). We will denote these cones as \( C^+_M \) and \( C^-_M = -C^+_M \).

To produce a partial order from the quadratic form \( q_M \) we will need to choose one of the cones.
Definition 39. Let $M$ be a Minkowski space whose set of causal vectors is the union of the cones $C^+_M, C^-_M$ as above. A time orientation of the Minkowski space is the choice of one of these cones denoted $C_M$.

From now on we will assume that a Minkowski space $M$ comes equipped with a time-orientation. We can define a partial order on Minkowski space.

Definition 40. Let $M$ denote a Minkowski space with a time orientation $C_M$. The causal partial order $\leq_M$ on $M$ is defined by $v \leq_M w$ iff $q_M(v, w) = (w - v)M(w - v)^t \geq 0$ and $w - v \in C_M$. We also define the relation $v <_M w$ iff $v \leq_M w$ and $w - v$ is time-like.

From the definition it follows that reversing the time orientation, that is, choosing $-C_M$ instead of $C_M$ reverses the partial order. In addition we see that $v$ is a causal vector iff $v \leq_M 0$ or $0 \leq_M v$.

6.1.2. Examples of causal partial orders. When $n = 1$, it is clear that the causal partial order associated with the standard Minkowski space matrix $D_{1,1}$ (with coordinates $q, r$) and the time orientation $C^+$, is $\leq_{ver}$ which played an important role in the disk scheduling discussion.

The partial order $\leq_{hor}$ which is complementary to $\leq_{ver}$ is the causal partial order associated with the matrix $-D_{1,1}$. This is true more generally in the $n = 1$ case, where if $M$ is a Minkowski matrix so is $-M$ and the causal partial orders associated with $M$ and $-M$ are complementary.

Recall that $\leq_{ver}$ is mapped to $\leq_{pos}$ via a 45 degree clockwise rotation, $R$. If we use the coordinates $t, x$, the matrix $P = RD_{1,1}R^t$ has the form

\[
\begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}
\]

The associated quadratic form is

\[
u(RD_{n,1}R^t)u = 2tx
\]

and it is easy to see that $\leq_{pos}$ is indeed the causal partial order.

6.1.3. Symmetries of Minkowski space.

Definition 41. Let $M$ be a Minkowski space. A linear transformation given by a matrix $L$ such that $LML^t = M$ is called a Lorentz transformation.

Since these transformations preserve the quadratic form associated with $M$ they are the symmetries of the Minkowski space.

To see what Lorentz transformations look like we can check the case $n = 1$. Instead of considering the standard matrix $D_{1,1}$ it is easier to consider matrices $L$ which preserve the quadratic form associated with $M = RD_{1,1}R^t$ which was introduced above in relation to $\leq_{pos}$. Consider the diagonal matrices $D_a$ with entries $(a, a^{-1})$, then it is easy to check that $D_aMD_a^t = M$. If $a > 0$ then $D_a$ will also preserve the partial order, otherwise, if $a < 0$ it will reverse it since it switches the time orientation. Reflection along the diagonal
and anti-diagonal also preserve $M$ and these transformations together with the transformations $D_a$ generate all Lorentz transformations.

**Remark:** We note that the causal partial order $\leq_{pos}$ associated with $M$ is preserved in this case by a larger group than the quadratic form preserving transformations. It is preserved by the mappings of the form $(x, y) \rightarrow (f(x), g(y))$, where both $f$ and $g$ are monotone increasing. These are precisely the mappings that we used before to normalize the distribution in airplane boarding problem. For $n > 1$, by a theorem of Zeeman [192], any 1-1 and onto map (linear or not) of $\mathbb{R}^{n+1}$ to itself, which maps the origin to itself, and preserves the causal partial order corresponding to $M$, is up to scaling (multiplication by a scalar) a Lorentz transformation.

We can ask, to what extent does the partial order $\leq_M$ given by $M$ and a cone $C_M$, determine $M$ itself. Obviously if we take the matrix $cM$ with $c > 0$ and the cone $C_{cM} = C_M$ we get the same partial order. It can be shown that this is the only freedom we have, that is, $\leq_M$ determines $M$ up to scaling. For $n = 1$ this follows from the discussion above on transformations that preserve $\leq_{pos}$ and the general case follows from Zeeman’s theorem or some elementary vector geometry.

**Lemma 8.** Lorentz transformations act transitively on time-like vectors $v$ such that $q_M(v) = 1$.

**Proof:** For $n = 1$ consider the Minkowski matrix $\frac{1}{2}RD_1,1R^t$. By equation (26) for $v = (t, x)$ we have $q_M(v) = tx$. A vector $v$ is time-like iff $tx > 0$. We assume that both $x, t > 0$, otherwise we can apply to $v$ the linear transformation $-Id(v) = -v$, which is always Lorentzian. Applying $D_a$ to $v$ with $a = \sqrt{x/t}$ we obtain $D_a(v) = (\sqrt{tx}, \sqrt{tx}) = \sqrt{q_M(v)}(1, 1)$. If $v, w$ are time-like with $q_M(v) = Q_M(w)$, they can be mapped to the same vector and hence can be mapped to each other, proving the assertion in this case.

Since all Minkowski matrices of the same dimension differ only by invertible linear transformations, proving the assertion for one proves it for all in the same dimension, hence the case $n = 1$ is proved. consider now the standard Minkowski matrix $D_{n,1}$ for general $n$. It is easy to see that any linear transformation which fixes the $t$ coordinate line and acts as an orthogonal transformation on the space defined by $t = 0$ is Lorentzian. Given a time-like point $(t, x_1, ..., x_n)$ we can rotate it using such a transformation into a point of the form $(t, x, 0, ..., 0)$, since orthogonal transformations can map any point $(x_1, ..., x_n)$ into any other point on the sphere of the same radius. We can then use a Lorentzian transformation on the $(t, x, 0, ..., 0)$ plane which acts as the identity on the complement given by, $t = x_1 = 0$, to map the point to $\sqrt{q_M(v)}(1, 0, ..., 0)$. q.e.d.
As a consequence of the lemma, we can map any pair of points \( u, v \) such that \( u - v \) is time-like into the pair \( 0, (1, 0, \ldots, 0) \) using a translation \( T(u) = u + z \) for some fixed \( z \), followed by a Lorentz transformation, followed by a scaling transformation \( S(v) = cv \) for some constant \( c > 0 \). The same consequence is true, with essentially the same argument, for all pairs \( u, v \) such that \( u - v \) is light-like and for all pairs such that \( u - v \) is space-like.

6.1.4. Intervals. We make several general definitions for posets that we will later apply to the causal partial order.

**Definition 42.** Given a point \( A \) in a poset \( M \) with partial order relation \( \leq_M \), we define the **past of** \( A \) to be the set
\[
I^-_M(A) = \{ B \mid B \leq_M A \}
\]
and the **future of** \( A \)
\[
I^+_M(A) = \{ B \mid A \leq_M B \}
\]
Given two points \( A, B \in M \) such that \( A \leq_M B \) we define the (causal) interval between them as
\[
I_M(A, B) = I^-_M(B) \cap I^+_M(A)
\]
Equivalently, this is the set of points \( C \) which satisfy \( A \leq_M C \leq_M B \).

If \( M \) is a Minkowski space and \( A <_M B \), then, we define the **open interval** \( I^o_M(A, B) \) to be the set of points \( C \in M \) such that \( A <_M C <_M B \).

Geometrically, for Minkowski space, the past and future of \( v \) are the translated cones \( C_M + v \) and \( -C_M + v \). The interval for points \( A, B \) such that \( B - A \) is time-like, is the intersection of two cones in opposing directions. If \( B - A \) is light-like it is easy to verify that \( I_M(A, B) \) consists of the line segment from \( A \) to \( B \). This is best seen by considering the case \( n = 1 \) and \( \leq_{\text{pos}} \).

**Definition 43.** A set \( D \subset M \) is said to be **causally convex** if for any \( A, B \in D \) such that \( A \leq_M B \) we have \( I_M(A, B) \subset D \).

It is clear from the definition that for any \( A \leq_M B \), the set \( I_M(A, B) \) itself is causally convex and that the same holds for \( I^o_M(A, B) \) when \( A <_M B \).

6.1.5. Causal curves. We can use the partial order to define curves which are chains w.r.t. the partial order. More explicitly.

**Definition 44.** Let \( M \) be a Minkowski space. A continuous map \( \gamma : [a, b] \to M \) from a closed interval to \( M \) with the standard topology is called a **parametrized curve**. We call the image in \( M \) of such a map a **curve**. We say that the parametrized curve is a curve from \( A = \gamma(a) \) to \( B = \gamma(b) \). A parametrized curve is said to be **order preserving** if it satisfies \( \gamma(u) \leq_M \gamma(v) \) whenever \( u \leq v \). We say that a curve is **causal** if it is the image of an order preserving map. Equivalently, a causal curve is a chain in \( M \) which is the continuous image of a closed interval.
If $A$ and $B$ are points such that $A \preceq_M B$ then the line segment from $A$ to $B$ is easily seen to be a causal curve from $A$ to $B$. There are simple parameters for causal curves in Minkowski space. Let $M$ be a Minkowski matrix. Let $v$ be a vector which is an eigenvector of $M$ with positive eigenvalue and $v \in C_M$. Let $L_v$ be the line passing through $v$ and let $T : \mathbb{R}^{n+1} \to L_v$ be the standard orthogonal projection onto the line given by $T(w) = \frac{wv'}{\sqrt{vv't}}v$. Assume we consider a basis consisting of eigenvectors of $M$, of unit Euclidean length. Since $M$ is symmetric the eigenvectors will be orthogonal and the time coordinate of a point w.r.t. this basis will be given by the orthogonal projection onto $L_v$. This leads to the following definition

**Definition 45.** A time parameter on Minkowski space $M$ is a function of the form $c + \frac{ww't}{\sqrt{vv't}}$ on points $w \in \mathbb{R}^{n+1}$ where $c$ is a constant and $v \in C_M$ is an eigenvector of $M$.

When the Minkowski metric is $D_{n,1}$, the $t$ coordinate of a vector is a time parameter. Given a causal curve, we can think of it as a function from the real line, given by the $t$ coordinate as the parameter to $\mathbb{R}^n$ given by the $x_1, \ldots, x_n$ coordinates. Viewed this way the notion of a causal curve coincides with that of a Lifshitz function with Lifshitz constant 1. We can also think of the curve as a function from the real line with the $t$ coordinate to Minkowski space $\mathbb{R}^{n+1}$, mapping $t$ to $(t, x_1(t), \ldots, x_n(t))$. Any causal curve will still be Lifshitz when $\mathbb{R}^{n+1}$ is given any Euclidean metric. This is the way we will describe causal curves as being Lifshitz in what follows. For a general Minkowski metric $M$ the same arguments apply to show that causal curves parametrized by a time parameter of $M$, are also Lifshitz functions w.r.t. the standard metric on $\mathbb{R}^{n+1}$, perhaps with a different constant.

Since a causal curve is Lifshitz the causal curve always has finite length w.r.t. any Euclidean metric.

6.1.6. Lorentzian length of a causal curve. We can use the Minkowski metric to define the length of causal curves in Minkowski space, much in the same way that the usual Euclidean/Pythagorian quadratic form $\sqrt{\sum_{i=1}^n x_i^2}$ is used to define the length of curves in Euclidean space. We begin by considering the analogue of distance

**Definition 46.** Given points $A, B \in \mathbb{R}^{n+1}$ such that $v = B - A$ is a causal vector, we define the **Minkowski distance** between $A$ and $B$ to be

$$d_M(A, B) = \sqrt{vMv^t} = \sqrt{q_M(B - A)}$$

Given a causal curve $\gamma : [a, b] \mapsto M$ from $A$ to $B$ we can use the quadratic form to assign a length to it. We follow the same procedure that is used to define length in Euclidean space. As usual we will consider the case of the standard quadratic form. We consider an ordered set of points on a causal curve $\gamma$ given as $A = E_1 \leq E_2 \leq \ldots \leq E_n = B$. We sum the Minkowski
distances between consecutive points

\[ \sum_{i=1}^{n-1} d_M(E_i, E_{i+1}) \]

to obtain an estimate on the length of the curve. In the Euclidean case we would do the same with \( d_K \) instead of \( d_M \).

If the curve would have been linear between the points then that would be its actual length. We make better and better approximations by adding more points to the computation. When we do that with a Euclidean metric, the usual triangle inequality makes the estimates non-decreasing and we can define the length of the curve to be the supremum, over all finite sets of points, of the length estimate. In Minkowski space the effect of adding points is the opposite! As we add points, the length estimate is non-increasing. To see this, we work with the standard Minkowski matrix \( M = D_{n,1} \). Suppose we added a point \( E \) on the curve, between \( E_i \) and \( E_{i+1} \). If \( E_{i+1} - E_i \) is light-like then \( d_M(E_{i+1}, E_i) = 0 \) and \( I_M(E_i, E_{i+1}) \) is the line segment between them. Since we must have \( E \in I_M(E_i, E_{i+1}) \) and for any such \( E \), \( d_M(E_i, E) = d_M(E, E_{i+1}) = 0 \), the claim is trivially true. If \( E_i <_M E_{i+1} \), we can assume via transitivity of the Lorentz transformations on such pairs that \( E_i \) is the origin and that \( E_{i+1} = (d_M(E_i, E_{i+1}), 0, ..., 0) \). Looking at the argument for transitivity, we see that we can also assume that \( E \) has the form \((t, x_1, 0, ..., 0)\). We have

\[ d_M(E_i, E) = \sqrt{t^2 - x_1^2} \leq t \]

and

\[ d_M(E, E_{i+1}) = \sqrt{(d_M(E_i, E_{i+1}) - t)^2 - x_1^2} \leq d_M(E_i, E_{i+1}) - t \]

which proves the assertion by adding the two inequalities. We also note that equality holds only if \( E \) is on the line segment between \( E_i \) and \( E_{i+1} \).

Following the above observations, we make the following definition,

**Definition 47.** Let \( \gamma \subset M \) be a causal curve. The **length** of \( \gamma \) denoted as \( \ell(\gamma) = \ell_M(\gamma) \), is the infimum of the expression (27) over all sequences of events \( A = E_1 \leq E_2 \leq \ldots \leq E_n = B \) on the curve.

Note that the length does not depend on a particular parametrization of the curve and that the length of the line segment from \( A \) to \( B \) is \( d_M(A, B) \).

An important corollary of the computation above is that for any causal curve \( \gamma \) starting at a point \( A \) and ending at \( B \), we have \( \ell(\gamma) \leq d_M(A, B) \), and equality holds only for the straight line between \( A \) and \( B \). Thus, the straight line between the points maximizes the length among all causal curves starting at \( A \) and ending at \( B \). In fact, this argument shows that for any causal curve \( \gamma \subset I(A, B) \) we have \( \ell(\gamma) \leq d_M(A, B) \).

The following lemma provides a more computable version of the length.
Lemma 9. Consider a twice differentiable parametrized causal curve, \( \gamma(s) \) with endpoints \( \gamma(a) = A, \gamma(b) = B \). Let \( v(s) = \gamma'(s) \) be the tangent to the curve at \( \gamma(s) \), given by differentiating each coordinate separately. Then, the length of the curve \( \ell_M(\gamma) \) is given by

\[
\ell_M(\gamma) = \int_a^b \sqrt{v(s)Mv(s)^t} \, ds
\]

Like all other results in this section, the claim is proved in the appendix. In fact, it is shown that the formula, appropriately interpreted, holds for all causal curves without the differentiability assumption.

The following theorem states an important property of the length function on the space of causal curves.

Theorem 15. Let \( \gamma \) be a causal curve and let \( \varepsilon > 0 \) be some constant. There exists an open set \( U = U(\gamma, \varepsilon) \) such that \( \gamma \subset U \) and such that for any causal curve \( \beta \subset U \) we have \( \ell(\beta) < \ell(\gamma) + \varepsilon \).

6.1.7. Orthogonality and volume. Apart from “lengths” the quadratic form associated with Minkowski space, provides us with notions of orthogonality and volume.

Definition 48. Let \( M \) be a symmetric matrix. We say that \( v \) and \( w \) are \( M \)-orthogonal if \( vMw^t = 0 \).

If \( D \) is diagonal then the standard basis vectors \( e_i \) are \( D \)-orthogonal.

Definition 49. Let \( M \) be a Minkowski metric and \( D \) a measurable set of \( \mathbb{R}^{n+1} \) whose usual Lebesgue/Euclidean volume is denoted by \( Vol_{\text{euc}}(D) \). We define the Minkowski volume of \( D \) to be

\[
Vol_M(D) = \sqrt{|\text{Det}(M)|Vol_{\text{euc}}(D)}
\]

Note that if \( M = TD_n,1T^t \) we have \( |\text{det}(M)| = |\text{det}(T)|^2|\text{det}(D_n,1)| \).

From this and the way the Euclidean volume changes under a linear change of coordinates, we can see that the Minkowski volume is compatible with linear changes of coordinates, i.e.,

\[
Vol_{TM^nT^t} = |\text{det}(T)|Vol_M(D) = Vol_{M}(T(D))
\]

We noted before that in the case of a Minkowski space metric \( M \), the causal partial order determines \( M \) up to a positive scaling factor. The Minkowski volume of the unit cube in \( \mathbb{R}^{n+1} \) can be used to determine the constant, thus the volume form and causal partial order fix the Minkowski metric completely.

For a causal interval \( I_M(A, B) \), with \( A \leq_M B \) in a \( d = n + 1 \) dimensional Minkowski space we have the relation \( d_M(A, B) = c_d Vol_M(I_M(A, B))^{1/d} \), where \( c_d > 0 \) is some constant that depends only on the dimension. This follows from the compatibility of the volume w.r.t. linear transformations, the transitivity property of Lorentz transformations and the fact that both sides scale linearly.
6.2. General (Lorentzian) space-time manifolds.

6.2.1. The underlying geometry of a space-time. We would like to extend the notion of a Minkowski space to cover more general geometric objects. Minkowski space which has been described using a coordinate system on $\mathbb{R}^d$. However, most geometrical objects (manifolds) like the sphere or a bagel, for example, cannot be described using a single coordinate system. Instead, to describe such objects, one considers an open covering of the object by patches, also known as charts, which look like open sets of Euclidean space. This is analogous to an atlas, where in each page we see a different piece of the earth described in a certain coordinate system. We also want to allow the Minkowski metric which is used for computing lengths of curves to smoothly vary from one point to another. Such a varying metric is called a Lorentzian metric. The end result, a manifold with a Lorentzian metric will be called a space-time or a Lorentzian manifold.

We recall the definition of a manifold.

**Definition 50.** Let $S$ be a metric space with metric $d_S$. A **chart** is a map $\phi : U \to \mathbb{R}^d$, where $U \subset S$ is open and $\phi$ is a homeomorphism from $U$ onto its image and the image is homeomorphic to $\mathbb{R}^d$. Since we are mapping $U$ to $\mathbb{R}^d$ a chart assigns coordinates to any point in $U$. An **atlas** is a (countable) collection of charts $\phi_i : U_i \to \mathbb{R}^d$, with some fixed value of $d$ and where the $U_i$ form a covering of $S$, namely, $S = \bigcup_i U_i$.

Let $\phi_i, \phi_j$ be a pair of charts in an atlas. Let $V = \phi_i(U_i \cap U_j) \subset \mathbb{R}^d$ and $W = \phi_j(U_i \cap U_j) \subset \mathbb{R}^d$. We obtain maps $\phi_j \phi_i^{-1} : V \to W$ and $\phi_i \phi_j^{-1} : W \to V$.

**Definition 51.** We say that a metric space $S$ is a **smooth manifold of dimension** $d$, if $S$ has an atlas such that for any $i, j$ the maps $\phi_j \phi_i^{-1}$ and $\phi_i \phi_j^{-1}$ are infinitely differentiable, i.e., the transitions between coordinate systems are smooth maps.

6.2.2. Lorentzian metrics. So far we have described the underlying topological structure of a space-time. We need to also generalize the notion of a Minkowski metric. We first describe the case of a single coordinate system (chart).

We first motivate the definition. Assume that we have a coordinate system with $d = n + 1$ coordinates $(x_0, \ldots, x_n)$. Let $D \subset \mathbb{R}^d$ be some domain (nice subset) of the coordinate space. We want to attach to each point $x \in D$ a Minkowski metric, i.e., a non-singular symmetric matrix with a single positive eigenvalue and $n$ negative eigenvalues. Instead of $M$ we will use the notation $g_x$ to describe these matrices. The entries of the matrix must be smooth functions of the coordinates.

Minkowski space, that we have encountered before, corresponds to the special case where the matrix $g = g_x$ is fixed. Using the quadratic forms associated with these matrices we are going to generalize the notions of causality and of curve length. As in Minkowski space, this will be done by
considering expressions of the form

(29) \( \gamma' g \gamma'' \)

When defining the metrics on different charts we will need to insure that the computed lengths of the same curve which is described (in different coordinate systems) in two charts coincide, i.e., that causality and lengths of curves are independent of the coordinate system used in describing the curve. For that we need to understand how the entries of the matrix \( g \) will transform with the coordinate change to keep expressions such as (29) unchanged.

Suppose that we have a smooth invertible transformation of coordinates from a coordinate system \( x_0, ..., x_n \) to a coordinate system \( y_0, ..., y_n \). We can write the original coordinates of a point \( x \) in terms of the new coordinates \( x_0 = x_0(y_0, ..., y_n), ..., x_n = x_n(y_0, ..., y_n) \). Suppose we have a curve \( \gamma(s) = (x_0(s), ..., x_n(s)) \). After the change of coordinates the curve will be described as \( \gamma(s) = (y_0(s), ..., y_n(s)) \). We have

\[
\frac{d\gamma(s)}{ds} = \left( \frac{dx_0}{ds}, ..., \frac{dx_n}{ds} \right)
\]

The change of variable formula for the tangent vector of a curve is

\[
\frac{dx_i}{ds} = \sum_k dx_i \frac{dy_k}{ds}
\]

and from this we see that

\[
v_x g_x v^t_x = \sum_i \sum_j (g_x)_{i,j} \frac{dx_i}{ds} \frac{dx_j}{ds}
\]

\[
= \sum_i \sum_j \sum_k \sum_l (g_x)_{i,j} \frac{dx_i}{dy_k} \frac{dx_j}{dy_l} \frac{dy_k}{ds} \frac{dy_l}{ds}
\]

\[
= \sum_k \sum_l (g_y)_{k,l} \frac{dy_k}{ds} \frac{dy_l}{ds} = v_y g_y v^t_y
\]

where

\[
(g_y)_{k,l} = \sum_i \sum_j (g_x)_{i,j} \frac{dx_i}{dy_k} \frac{dx_j}{dy_l}
\]

This tells us how the matrices should transform under a change of coordinates, to preserve quantities of interest. The discussion leads to the following definition

**Definition 52.** A **Lorentzian metric** on a smooth manifold \( S \), or equivalently, a space-time structure on \( S \) is defined by the choice of a point dependent Minkowski metric on \( \phi_i(U_i) \) for each \( i \), where matrix entries are smooth functions of the coordinates. In addition, given a coordinate transformation \( \phi^{-1}_j \phi_i : U_i \rightarrow U_j \) mapping \( x \in U_i \) to \( y \in U_j \) (and vice versa), the Minkowski metrics \( g_x \) and \( g_y \) defined for the charts \( U_i \) and \( U_j \) are related
A smooth manifold \( S \) equipped with a Lorentzian metric is known as a \textbf{Lorentzian manifold} or a \textbf{space-time}.

As an example, consider an infinite cylinder \( C \) with coordinates \((r, \theta)\), with the identification of \( \theta \) with \( \theta + 1 \). We recall that in our modeling, a disk drive was identified with a compact piece of this cylinder consisting of points with \( 0 \leq r \leq 1 \). To view the cylinder as a manifold we consider the subsets \( A, B, C \) consisting of points with \( 0 < \theta < 2/3 \), \( 1/3 < \theta < 1 \) and \( 2/3 < \theta < 4/3 \) respectively. As atlases we consider the identity map of these sets into the \((r, t)\) plane, where the \( t \) coordinate coincides with the \( \theta \) coordinate but has no identifications. Since there are no identifications of the \( \theta \) coordinate in \( A, B \) or \( C \) the atlas condition of being a homeomorphism onto an open set is satisfied, and the transition functions are translations which are certainly smooth, so we have a manifold. At each point in each atlas we can take a Lorentz metric with associated quadratic form \( p(r, t)(t^2 - r^2) \), where \( p(r, t) \) is a smooth function which satisfies \( p(r, t) = p(r, t + 1) \). Since the transition functions are translations it is trivial to check the transformation rule, which yields the identity transformation. We will use the space-time manifolds of this form to analyze the disk scheduling problem. Another choice which plays a role in disk scheduling is to take the Lorentz metric with quadratic form \( p(r, t)(r^2 - t^2) \).

6.2.3. \textbf{Vectors}. We define the notion of a tangent vector attached to each point.

\textbf{Definition 53}. A (tangent) \textbf{vector} at the point \( s \in S \) is the choice of a vector centered at \( \phi_i(s) \) for each \( i \) such that \( s \in U_i \). In addition, if \( s \in U_i \cap U_j \), \( x = \phi_i(s) \), \( y = \phi_j(s) \), the coordinates of the assigned vectors \( v(x) \) and \( v(y) \) satisfy

\begin{equation}
\sum_{k} \frac{dx_i}{dy_k} v(y)_k
\end{equation}

\( v(x)_i = \sum_k \frac{dx_i}{dy_k} v(y)_k \)

A vector \( v(s) \) is \textbf{causal}, \textbf{time-like}, \textbf{light-like} or \textbf{space-like}, if the vector \( v(x) \) representing it in some chart with \( \phi_i(s) = x \) is causal, time-like, light-like or space-like, respectively, w.r.t. the Minkowski metric \( g_x \). The considerations above show that this is well defined.

We also need the notion of a smoothly varying vector.

\textbf{Definition 54}. A \textbf{vector field} is an assignment of a vector \( v(s) \) for all \( s \in S \), such that for each coordinate chart the coordinates of \( v(x) \) are smooth functions of the coordinates on \( \phi_i(U_i) \). A vector field is \textbf{non-vanishing} if \( v(s) \) is never the zero vector.

Using these definition we can define the analogue of time orientation for space-time manifolds.

\textbf{Definition 55}. A \textbf{time orientation} of a space-time is a non-vanishing vector field \( v(s) \) such that for each \( s \in S \), \( v(s) \) is causal.
The vector $v(s)$ chooses at each point $s$ one of the two cones of causal vectors, the one it belongs to. Since $v(s) \neq 0$, only one cone is chosen, and the choice varies smoothly, hence continuously.

6.2.4. Causal curves in Lorentzian geometry.

**Definition 56.** A differentiable parametrized curve $\gamma : [a, b] \to D$ is said to be causal if at any point $\gamma(s)$ we have that $\gamma'(s)$ is a causal vector.

The causal relation $R_S$ on a space-time $S$ with a time orientation $v(s)$ is defined by $(A, B) \in R_S$ iff there is a piecewise differentiable causal curve $\gamma(s)$ beginning at $A$, ending at $B$, and such that at each point $s$, the tangent $\gamma'(s)$ is a non-zero causal vector in the same cone $v(s)$.

We can remove the word piecewise from the definition, see [142] proposition 2.23. It follows from the definition that the relation is transitive.

The relation $R_S$ is not always anti-symmetric. Consider the cylinder $C$ with the Lorentz metric $t^2 - r^2$ that was introduced above. This space-time contains causal curves whose endpoints $A, B$ coincide, for example $\gamma(s) = (r, s)$ for some fixed $r$ and $0 \leq s \leq 1$. In such cases the relation $R_S$ will not give us a partial order.

**Definition 57.** A causal curve is said to be a closed causal curve if its endpoints $A, B$ coincide. A space-time $S$ with a time orientation is said to be a causal space-time if there are no closed causal curves in $S$. For a causal space-time, the causal relation $R_S$ is a partial order relation called the causal partial order.

A space-time $S$ is strongly causal if for each point $s \in S$ and any open set $U = U_s$ which contains $s$, there is an open set $N = N_{s,U}$ which satisfies $s \in N \subset U$ and such that for any parametrized causal curve $\gamma : [a, b] \to S$ we have $\gamma^{-1}(N)$ is an open interval in $[a, b]$. In particular, if $N$ contains the endpoints of $\gamma$ it contains the entire curve. An open set $N$ having the last property is called causally convex.

Strong causality means that non-constant causal curves cannot come back arbitrarily close to their initial point $A$. Some “ugly” examples, [142], show that it is a strictly stronger condition than causality, justifying the name.

6.2.5. Local approximation by Minkowski spaces.

**Definition 58.** Let $s \in S$ be a point in a space-time and let $U_s$ be a chart such that $s \in U_s$ equipped with a Euclidean metric $K$. Let $x = \phi_1(s)$. An $\varepsilon$-normal neighborhood of $s$ (w.r.t. $U_s$ and $K$) is an open set $N_{x,\varepsilon}$, such that

1) $x \in N_{x,\varepsilon}$.

2) There are Minkowski metrics $M_1, M_2$ on $\phi_1(U_s)$ such that

$C(M_1) \subset C(g_x) \subset C(M_2)$

holds on $N_{x,\varepsilon}$. The metric $M_1$ will be called the majorizing Minkowski metric while $M_2$ will be called the minorizing Minkowski metric.
3) $N_{x,\varepsilon}$ is causally convex w.r.t. $M_2$ and so w.r.t. the Lorentz metric on $S$ as well.
4) For any non-zero causal vector w.r.t. $M_2$ and for any $y \in N_{x,\varepsilon}$ either (213) or
\[ \sqrt{vM_2v^t} \geq \sqrt{vg_yv^t} \geq \sqrt{vM_1v^t} \geq \sqrt{vM_2v^t} \frac{1}{1-\varepsilon} \]
holds.

An open set satisfying only the first three conditions is called a normal neighborhood of $x$.

The following lemma is proved in the appendix.

**Lemma 10.** Let $s \in U_i$ be a point in a chart of a strongly causal space-time and let $K$ be a Euclidean metric on the chart. Then, for any $\varepsilon > 0$ there exists an $\varepsilon$-normal neighborhood of $s$ in $U_i$.

6.2.6. Lengths and volumes.

**Definition 59.** Consider a twice differentiable parametrized causal curve $\gamma(s)$ with endpoints $\gamma(a) = A$, $\gamma(b) = B$. Let $x = \gamma(s)$. Let $v_x$ be the tangent to the curve at $x$, given locally in a given coordinate chart by differentiating each coordinate separately. The length of the curve $\ell_M(\gamma)$ is given by
\[ \ell_M(\gamma) = \int_a^b \sqrt{v_xM_xv^t} ds \]

We note that our definition of a Lorentzian metric ensures us that the formula is well defined regardless of the local coordinate system which we use at any given point.

We can also use normal neighborhoods to define the length of causal curves in $S$. Let $\gamma$ be a parametrized curve in $S$ with endpoints $A$ and $B$. For any $s \in [a, b]$ let $N_s$ denote a normal neighborhood containing the point $\gamma(s)$. The sets $\gamma^{-1}(N_s)$ form an open covering of $[a, b]$ which is a compact set, hence we can find a finite sub-cover $N_i = N_{s_i}$, $i = 1, ..., k$. Each $N_i$ has a majorizing Minkowski metric $g_i$ such that $vg_xv^t < vg_i v^t$ for any $v$ centered at $x \in N_i$.

Let $A = E_0 \leq_S E_1 \leq_S ... \leq_S E_m = B$ be a sequence of points on the curve such that for any $0 \leq j \leq m - 1$ there is an $i(j)$ such that $E_j, E_{j+1} \in \phi(N_{i(j)})$. We let $\ell(N_i, E_j) = d_{i(j)}(E_j, E_{j+1})$, where $d_{i(j)}$ refers to the Minkowski distance function w.r.t. the Minkowski metric $g_i$ on $\phi_i(N_i)$.

**Definition 60.** The length $\ell(\gamma)$ of a causal curve $\gamma$ is defined as the infimum of the expressions $\ell(N_i, E_j)$, over all possible $N_i, E_j$ which satisfy the assumptions above.

Using lemma 9 and the local approximation using Minkowski metrics it is easy to check that this definition coincides with the definition using formula (36).
The volume form in a general space-time is also defined as before. For the vectors $e_0, e_1, \ldots, e_n$ in a coordinate system at a point $x$, it gives the volume $\sqrt{|\text{Det}(g_x)|}$. The change of coordinates formulas for the tangent vectors $e_0, e_1, \ldots, e_n$ and for $g_x$ can be used to verify that this is well defined regardless of the particular chart used.

6.2.7. The topology of the space of causal curves. Let $S$ be a space-time and let $D \subset S$ be a compact subset of $S$. Let $A, B \subset D$ be a pair of compact subsets of $D$ and consider $C_D(A, B)$, the set of causal curves in $D$ starting at some $a \in A$ and ending at some $b \in B$.

We define a topology on this set.

**Definition 61.** If $N$ is an open subset of $D$ let $C_N(A, B)$ be the set of causal curves in $N$ starting in $A$ and ending in $B$. The $C_0$ topology on $C_D(A, B)$ is the topology whose basis is the collection of open sets $C_N(A, B)$.

In addition to this topology we also have the Hausdorff metric defined on $C_D(A, B)$, where causal curves are considered as closed subsets of the metric space $S$.

The following result is proved in [142], theorem 6.5. The proof there is a bit sketchy, so for the convenience of the reader we provide essentially the same proof but with added details in the appendix.

**Theorem 16.** Assume that the space-time $S$ is strongly causal and let $A, B \subset D$ be compact sets in $S$. Then, the set $C_D(A, B)$ is compact in the Hausdorff metric and in the $C_0$ topology.

Theorem 15 states that the length function on causal curves in Minkowski space is upper semi-continuous in the $C_0$ topology. The following result states the same for strongly causal space-time manifolds $S$, the proof is also essentially the same.

**Theorem 17.** Let $S$ be a strongly causal space-time. The length function is upper semi-continuous on causal curves in $S$ with the $C_0$ topology.

Since an upper semi-continuous function on a compact space attains a maximum we obtain the following result.

**Theorem 18.** Let $S$ be a strongly causal space-time and let $A, B \subset D$ be compact subsets of $S$, then, there are curves of maximal length in $C_D(A, B)$.

As a corollary, if we set $A = B = D$ we see that there is a curve of maximal length in $D$. We are led to the following definitions

**Definition 62.** Let $D$ be some compact domain in a space-time $S$. A causal curve in $D$ of maximal length is called a maximal curve. The length of a maximal curve in $D$ will be called the diameter of $D$ and we will denote it by $\text{diam}(D)$.
7. Discrete space-time

Space-time is an object with a continuous geometry, however, our intended applications which are to disk scheduling and airplane boarding are discrete in nature. To bridge this gap we would like to consider discrete analogues of space-time.

**Definition 63.** Let $D$ be a compact subset of a strongly causal space-time $S$. Let $\mu_D = \text{vol}_{S}/\text{vol}_{S}(D)$ be the volume form on $S$ normalized so that $\mu(D) = 1$. The **discrete space-time** on $n$ points associated with $D$ is the random partially ordered set $X$ obtained by sampling $n$ points in $D$ w.r.t. $\mu_D$, equipped with the causal order on $S$ restricted to the sampled points. A discrete space-time is also called a **causal set**. We let $L_n(D)$ denote the random variable given by length of the longest chain in $X$.

Consider the cylinder $C$ with Lorentz metric $p(r, \theta)D_{1,1}$. Let $D$ be the compact domain consisting of points $(r, \theta)$ with $0 \leq r \leq 1$. The service time random variable $L_{n,p}$ that was introduced previously coincides with $L_n(D)$. We see that the analysis of the service time in the disk scheduling problem then becomes a special case of the analysis of $L_n(D)$ for a general compact domain in a space-time.

Discrete space-time was originally introduced in physics papers of 't Hooft and Myrheim, [167, 138] in an attempt to reconcile relativity theory and quantum mechanics. In this approach the universe is actually a discrete partially ordered set with about $10^{160}$ points and the notion of space-time as we have considered it is just a convenient approximation to the actual discrete object. Later, this approach led to the causal set program for quantum gravity, championed by R. Sorkin and his collaborators, [38]. This rather minimalistic approach is still actively pursued.

The following result, which is a generalization and re-interpretation of a result of Deuschel and Zeitouni, [61], shows that $L_n(D)$ is the discrete analogue of the diameter of $D$. This is expected since the number of points in a chain should be the discrete counterpart to the length of a chain (causal curve) in the continuous setting. We need a definition first.

**Definition 64.** We say that a compact subset (domain) $D$ of a strongly causal space-time $S$ is a **nice domain** if it satisfies the following two properties

(i) If two points $A, B \in D$ satisfy $A \leq_S B$ then there is a piecewise causal curve in $D$ that starts in $A$ and ends in $B$. Stated otherwise, when restricted to $D$, the notion of causality is the same as in $S$.

(ii) For any $\varepsilon > 0$, there is a causal curve $\gamma$ which is contained in the interior of $D$, such that $\ell(\gamma) > \text{diam}(D) - \varepsilon$.

**Theorem 19.** Let $D$ be a nice domain of unit volume. Let $\text{diam}(D)$ denote the maximal length of a causal curve in $D$. Let $L_n(D)$ be the maximal
length of a chain among \( n \) points, sampled according to the volume distribution on \( D \). Let \( X_n(D) \) be the total number of chains of all sizes among \( n \) sampled points. Then,

1) For all \( \varepsilon > 0 \), w.h.p., we have
\[
|L_n(D) - c_d \text{diam}(D)n^{1/d}| < \varepsilon n^{1/d}
\]

Moreover, w.h.p., there is a maximal curve \( \gamma \) such that all points \( x \) in a maximal chain satisfy \( d_S(x, \gamma) < \varepsilon \), i.e., the maximal chain will be contained in an \( \varepsilon \) neighborhood of a maximal curve.

2) For all \( \varepsilon > 0 \), w.h.p., we have
\[
|\log(X_n(D)) - \rho_d \text{diam}(D)n^{1/d}| < \varepsilon n^{1/d}
\]
where \( \rho_d \) depends only on the dimension \( d \).

Moreover, w.h.p., there is a maximal curve \( \gamma \) such that all points \( x \) in a uniformly random chain satisfy \( d_S(x, \gamma) < \varepsilon \), i.e., w.h.p. any chain is contained in an \( \varepsilon \) neighborhood of a maximal curve.

When \( d = 2 \) the constant \( c_2 \) has been determined by a deep result of Vershik and Kerov, [179], and independently, Logan and Shepp, [123]. Their result allows us in the two dimensional case to make the following more precise statement.

**Theorem 20.** Assume \( d = 2 \). Let \( D \) be a nice domain of volume 2. Assume that for any \( \varepsilon > 0 \), there is a causal curve \( \gamma \) which is contained in the interior of \( D \), such that \( \ell(\gamma) > \text{diam}(D) - \varepsilon \). Then,

1) For all \( \varepsilon > 0 \), w.h.p., we have
\[
|L_n(D) - \text{diam}(D)n^{1/d}| < \varepsilon n^{1/d}
\]

The rest of the section is devoted to proving theorem 19.

### 7.1. Poisson sampling

Before we continue to examine causal sets we alter the sampling process a bit. The problem is that if we sample exactly \( n \) points from some domain \( D \), then the subsets of sampled points in disjoint sub-domains, say \( A, B \) will be correlated. Basically, if we have a sampled point \( p \in A \) then we are a little less likely to have a sampled point in \( B \). The trick to overcome this problem is called Poissonization. We recall the following definition

**Definition 65.** The Poisson distribution with parameter \( n \) has values \( m = 0, 1, 2, ... \) such that the probability of \( m \) is given by
\[
p_{m,n} = e^{-n} \frac{n^m}{m!}
\]

We will denote this distribution by \( P(n) \).
Applying the Poisson distribution to sampling from a space-time domain we obtain the following definition.

**Definition 66.** Let $D$ be a compact measurable subset of a space-time $S$. A **discrete space-time on $P(n)$ points** is the random variable partially ordered set obtained by first sampling a value $m$ using a Poisson distribution with average $n$ and then sampling $m$ points from $D$ using the normalized volume form $\mu_D$. We let $\tilde{L}_n(D)$ denote the random variable given by the length of a maximal chain in a discrete space-time on $P(n)$ points.

It turns out that using this trick, the number of points sampled from a sub-domain $A$ of normalized volume $v_A$ will be a Poisson random variable $X_A$ with average $v_A n$. Moreover, if $A, B$ are disjoint, then the random variable $X_A, X_B$ are independent. To see why, lets work backwards. Assume that $D$ is divided into a finite number of sets $D_i$, with volume $v_i$, which are either disjoint, or more generally, their intersections have measure zero. Assume we sample independently $P(v_i n)$ points from each $D_i$. What would the entire sample look like. For the Poisson distribution with average $n$, $P(n)$, we have the generating function

$$f_{P(n)} = \sum_m p_{m,n} z^m = e^{-n(nz)^m/m!} = e^{nz-n}$$

having sampled independently $P(v_i n)$ points from each $D_i$, the total sample on $D$ will have distribution $\sum_i P(v_i n)$ whose generating function is $\prod_i e^{v_i n(z-1)} = e^{n(z-1)}$ which is the generating function for $P(n)$. In addition, within each $D_i$ the points are distributed as if sampled independently from the volume form.

Our next goal is to show that the actual number of points sampled using $P(n)$ is rarely much different than $n$. Towards that end we will estimate for $\eta > 1/2$ the probability $Pr(|m-n| \geq n^\eta)$ as $n \to \infty$. Notice that $p_{m,n}$ is an increasing function of $m$ for $m \leq n$ and a decreasing function when $m \geq n$. For $m \geq n + n^{1/2}$ we have

$$p_{m,n}/p_{m+1,n} = \frac{m+1}{n} \geq \frac{n + n^{1/2}}{n} = 1 + \frac{1}{\sqrt{n}}$$

Iterating we get for $m \geq n + n^\eta$

$$p_{n+n^{1/2},n}/p_{m,n} \geq (1 + \frac{1}{\sqrt{n}})^{n^\eta-n^{1/2}} = ((1 + \frac{1}{\sqrt{n}})^{\sqrt{n}})^{n^{\eta-1/2}-1}$$

Letting $n \to \infty$ and defining $\delta = \eta - 1/2$ we get that for any $n$ large enough and any $m \geq n + n^\eta$

$$p_{m,n} \leq 2^{-n^{\delta}}$$

where we have used the facts that $2 < e$ and $p_{n+n^{1/2},n} \leq 1$. Also, if $m \geq 2n$ then $p_{m,n}$ decreases faster than a geometric series with ratio 1/2. By summing separately $n + n^{1/2} \leq m < 2n$ and $m \geq 2n$ we obtain the estimate

(40) \quad \text{Pr}(m \geq n + n^\eta) \leq n2^{-n^{\delta}}
for all \( n \) large enough. A similar argument applies to \( Pr(m \leq n - n^\eta) \). We could provide more exact estimates but these will do for our purposes.

Poisson sampling has been used in the physics literature alluded to above.

### 7.2. Chains in discrete space-time

We will think of the discrete space-time on \( n \) or \( P(n) \) points as an approximation to the compact space-time domain \( D \). As the number of sampled points increases, we expect the resulting discrete set to provide an increasingly good approximation to \( D \), the continuous space from which it was sampled. Thus we can expect that as \( n \) becomes large, with high probability, the statistical properties of the discrete version will be given in terms of the original space-time geometry of \( D \). In particular, this should apply to the length of the longest chain.

Let us consider the case of an interval \( D = I(A,B) \) in Minkowski space, where \( A < M_B \). We may recall that in Minkowski space, as a partially ordered set, the interval is essentially unique, since we can use translations, dilations and Lorentz transformations, which all preserve order, to map any pair of points such that \( A < M_B \) to any other pair \( E, F \) satisfying the same condition. These maps also preserve the volume form up to proportionality and therefore identify the point sampling processes. Therefore, we will denote \( L_n(D) \) and \( \tilde{L}_n(D) \) in this special case simply by \( L_n \) and \( \tilde{L}_n \) without mention of the particular \( A, B \).

Understanding the behavior of \( L_n \) as \( n \) tends to infinity in the two dimensional case is known as Ulam’s problem. It has a long history. Stanislav Ulam used Monte Carlo (numerical) methods to study \( L_n \), [178]. His numerical studies have led him to assert that the value of \( L_n \) concentrates around \( \alpha \sqrt{n} \) for a specific value of \( \alpha \).

**Definition 67.** We say that a statement \( A_n \) is true with high probability, w.h.p., if the probability that the statement is false tends to 0 as \( n \) tends to infinity.

To be more precise, Ulam made the following conjecture, that was proved by Hammersley in [84].

**Theorem 21.** There is a constant \( \alpha > 0 \) such that for any \( \varepsilon > 0 \), w.h.p. \[ |L_n - \alpha \sqrt{n}| < \varepsilon \sqrt{n}. \]

We will prove a generalization of this result to nice compact domains \( D \) of a strongly causal space-time \( S \).

It is somewhat easier to analyze longest chains in intervals \( I(A,B) \) of a related partial order which generalizes the partial order \( \leq_{pos} \)

**Definition 68.** The positive partial order, \( \leq_{pos} \) on \( \mathbb{R}^d \) is defined by \( A \leq_{pos} B \) iff \( a_i \leq b_i \) for all \( i = 1, \ldots, d \). We let \( A <_{pos} B \) iff \( a_i < b_i \) for all \( i = 1, \ldots, d \).

Note that for \( d = 2 \) this coincides with our previous definition of \( \leq_{pos} \), but for \( d > 2 \) it no longer represents the partial order associated with a
Minkowski metric. Also note that for \(A <_{\text{pos}} B\) an interval \(I(A, B)\) is a \(d\) dimensional box. By stretching/contracting the coordinates separately we may assume without loss of generality that the box is the unit \(d\) dimensional cube which we denote by \(I^d\). This operation also preserves the standard Lebesgue measure up to a scaling constant, which we choose so that the unit cube has volume 1.

The next result, due to Bollobas and Winkler, \([37]\), provides asymptotic analysis for intervals in \(\leq_{\text{pos}}\) of any dimension. The technique is very similar to arguments of Hammersley and Kingman and can be considered an example of a sub-additivity argument, a rather common technique for proving concentration results. The positivity of the constant is based on an argument of Kingman, \([112]\). The proof technique also yields an estimate on the total number of chains which also holds w.h.p.

**Theorem 22.** Let \(I^d\) be the \(d\) dimensional unit cube in \(\mathbb{R}^d\). Let \(L_{n,\text{pos}}\), respectively \(\tilde{L}_{n,\text{pos}}\), denote the length of the longest chain w.r.t. \(\leq_{\text{pos}}\) among \(n\), respectively \(P(n)\), uniformly sampled points in \(I^d\). Let \(X_{n,\text{pos}}, \tilde{X}_{n,\text{pos}}\) denote the total number of chains among \(n, P(n)\) uniformly sampled points respectively. Then

1) There is a constant \(\alpha_d > 0\) such that for any \(\varepsilon > 0\), w.h.p.
\[
|L_{n,\text{pos}} - \alpha_d n^{1/d}| < \varepsilon n^{1/d}
\]
The same estimate with the same value of \(\alpha_d\) also holds for \(\tilde{L}_{n,\text{pos}}\).

2) There is a constant \(\beta_d > 0\) such that for any \(\varepsilon > 0\), w.h.p.
\[
(\beta_d - \varepsilon)^{n^{1/d}} < X_{n,\text{pos}} < (\beta_d + \varepsilon)^{n^{1/d}} < \varepsilon n^{1/d}
\]
The exact same estimate holds for \(\tilde{X}_{n,\text{pos}}\).

**Proof:** We follow the elementary proof of Bollobas and Winkler, \([37]\), but in a Poissonized version. We start with the first statement.

We let
\[
\alpha_d = \limsup_m E(L_{m,\text{pos}})/m^{1/d}
\]
We claim that \(\alpha_d\) is finite and satisfies the assertion of the theorem. We first show that \(\alpha_d \leq e\), where \(e\) is the natural logarithmic base. Consider a sequence of \(k\) points \(x_i = (x_{1, i}, ..., x_{d, i})\). We can sort these points in increasing order according to any of the \(d\) coordinates of the points. This gives us \(d\) permutations \(Q_j\), \(1 \leq j \leq d\), of the \(k\) points. The points form a chain w.r.t. \(\leq_{\text{pos}}\) if and only if all the permutations \(Q_j\) are the same. In a cube with the uniform distribution the different coordinates of a sampled point are independent variables and for a given coordinate all permutations are equally likely. Consequently, if we have \(k\) points, the probability that all their associated permutations are identical is \(1/(k!)^{d-1}\), since the first permutation is arbitrary and all others must match it. Let \(X_{m,k}\) be the number of chains of size \(k\). We see that the expected number of chains
of size $k$, $E(X_{m,k})$ is given by $B_{m,k}/(k!)^{d-1}$, where $B_{m,k} = \frac{m!}{(m-k)!k!}$ is the binomial coefficient, giving the number of sets of size $k$ among all $m$ sampled points. Let $k = bm^{1/d}$ for some fixed $b$. We assume for simplicity of notation that $k$ is an integer. We can compute the asymptotic value as $m \to \infty$ of $B_{m,k}/k!$ by using Stirling’s formula which we apply to all the factorials in $B_{m,k}/(k!)^{d-1} = \frac{m!}{(m-k)!k!}$. We have

\begin{equation}
E(X_{m,k}) = \frac{m!}{(m-k)!} \sim \sqrt{\frac{m}{m-k}} \left( \frac{m}{m-k} \right)^{m-k} (m/e)^k
\end{equation}

Also $\frac{m}{m-bm^{1/d}} = \frac{1}{1-bm^{-(d-1)/d}}$. Applying a Taylor series expansion we have

$$
\frac{1}{1-y} = 1 + y + y^2 + \ldots = (1+y)(1+y^2+y^4+\ldots) = (1+y)(1+\delta_y y^2)
$$

with $\delta_y \to 1$ as $y \to 0$. Letting $y = bm^{-(d-1)/d}$ we consider $m$ sufficiently large so that $y < 1$ and $1 \leq \delta_y \leq 2$. We get

$$
\left( \frac{m}{m-k} \right)^{m-k} = \left( \frac{m}{m-bm^{1/d}} \right)^{m-bm^{1/d}}
$$

$$
= \left( \frac{1}{1-y} \right)^{m-bm^{1/d}} < (1+y)^m (1+2bm^{-2(d-1)/d})^m
$$

$$
\leq (1+y)^m (1+2bm^{-1})^m
$$

since $m = y^{-1}bm^{1/d}$ we have

$$(1+y)^m = ((1+y)^{1/y})^{m^{1/d}b} ((1+2bm^{-1})^{m/(2b)})^{2b}
$$

$$
< e^{m^{1/d}b} e^{2b} = e^k e^{2b}
$$

where we have used the inequality $(1 + \frac{1}{x})^x < e$, for say $x \geq 1$. We conclude that

$$
\frac{m!}{(m-k)!} < \sqrt{\frac{m}{m-k}} m^k e^{2b}
$$

We also have

$$
k!^{-d} \sim (2\pi k)^{-d/2} \left( \frac{e}{k} \right)^{kd}
$$

Putting the two estimates together and considering $m$ large enough so that

$$
\sqrt{\frac{m}{m-k}} (2\pi k)^{-d/2} e^{2b} < 1
$$

yields

\begin{equation}
E(X_{m,k}) = \frac{m!}{(m-k)!} k!^d < \left( \frac{m}{k!d} \right)^{k} e^{kd} = \left( \frac{e}{b} \right)^{kd}
\end{equation}

We note that by Markov’s inequality we have $Pr(L_{m,pos} \geq k) \leq E(X_{m,k})$. We conclude that for $b > e$ we have

$$
Pr(L_{m,pos} \geq bm^{1/d}) < (e/b)^{bdm^{1/d}}
$$

Since the longest possible chain among $m$ sampled points has length at most $m$ we conclude from the above calculations that for any $b > e$ we have

$$
E(L_{m,pos})/m^{1/d} \leq bPr(L_{m,pos} \leq bm^{1/d}) + m^{(d-1)/d} Pr(L_{m,pos} > bm^{1/d})
$$
\[ \leq b + m^{(d-1)/d}\left(\frac{\epsilon}{b}\right)^k \rightarrow b \]

from which we conclude that \(\alpha_d \leq \epsilon\).

We claim that \(\alpha_d = \limsup_m E(\bar{L}_{n,\text{pos}})/n^{1/d}\) as well. We have

\[ E(\bar{L}_{n,\text{pos}})/n^{1/d} = \left(\sum_m p_m,n E(L_{m,\text{pos}})\right)/n^{1/d} \]

From our previous calculations on the Poisson distribution it is easy to verify that the contribution to the average of \(m\) such that \(|m - n| > n^{2/3}\) satisfies

\[ \sum_{|m-n|\geq n^{2/3}} p_{m,n} E(L_{m,\text{pos}}) \leq \sum_{|m-n|\geq n^{2/3}} p_{m,n} m \rightarrow 0 \]

On the other hand \(\sum_{|m-n|\leq n^{2/3}} p_{m,n} \rightarrow 1\) and for all these values of \(m\) we have \((m/n)^{1/d} \rightarrow 1\) with a convergence estimate that depends only on \(n\) since \(1 - n^{-1/3} < m/n < 1 + n^{-1/3}\) and hence is uniform for all \(m\) in the range of interest. For all such \(m\)

\[ \alpha_d = \limsup_m E(L_{m,\text{pos}})/m^{1/d} = \limsup_m E(L_{m,\text{pos}})/n^{1/d} \]

again in a uniform fashion which depends only on the behavior of the extreme cases \(m = n + n^{2/3}\) and \(n - n^{2/3}\). By summing we get the desired equality.

We now proceed to the second part of the argument which uses crucially the scale invariance of \(\leq_{\text{pos}}\). Given some \(\epsilon_1 > 0\), take \(n_0\) so that

\[ E(\bar{L}_{n_0,\text{pos}})/n_0^{1/d} \geq \alpha_d - \epsilon_1. \]

Let \(l\) be an integer. We can break up the unit cube \(I^d\) into \(l^d\) sub-cubes \(S(i_1,\ldots,i_d), 0 \leq i_j < l\), consisting of points \((x_1,\ldots,x_d)\) such that \(i_j/l \leq x_j < (i_j + 1)/l\). In particular we can consider the "diagonal" sub-cubes \(S_i = S(i,\ldots,i)\) consisting of points \((x_1,\ldots,x_d)\) such that \(i/l \leq x_j \leq (i + 1)/l\), for all \(j = 1,\ldots,d\). Consider \(l^d\) Poisson samples \(P(n_0)\), one for each sub-cube \(S(i_1,\ldots,i_d)\). Together they produce a \(P(l^d n_0)\) sample on the unit cube. We consider the largest chain in each diagonal sub-cube \(S_i\) and denote it by \(\bar{L}_{n_0,i}\). Since the sub-cube \(S_i\) is a scaled down (and translated) version of the unit cube \(I^d\), these random variables, say \(X_i\), are i.i.d. with distribution given by that of \(\bar{L}_{n_0,\text{pos}}\). We can concatenate the largest chain in each sub-cube to obtain a chain in the unit cube of size \(X_1 + \ldots + X_l\). We can now use the law of large numbers to claim that for any \(\epsilon_2,\epsilon_3 > 0\) if \(l\) is large enough the length of that chain will be at least \(lE(\bar{L}_{n_0,\text{pos}})(1-\epsilon_2)\) with probability at least \(1-\epsilon_3\), which means that for all \(n\) of the form \(n = n_0l^d\), for \(l\) large enough with probability at least \(1-\epsilon_3\),

\[ \bar{L}_{n,\text{pos}}/n^{1/d} = \bar{L}_{n_0,\text{pos}}/(n_0^{1/d}) \]

\[ \geq (1-\epsilon_2)E(\bar{L}_{n_0,\text{pos}})/n_0^{1/d} \geq (1-\epsilon_2)(\alpha_d - \epsilon_1) \]

For any \(n\), we can choose \(l\) such that \(l^d n_0 \leq n \leq (l + 1)^d n_0\) and the result will follow by applying it to the left and right values of the inequality since their ratio tends to 1. To see that w.h.p. \(\bar{L}_{n,\text{pos}} < (\alpha_d + \delta)n^{1/d}\) for any fixed \(\delta > 0\), we note that by definition, for any \(j\) and for large enough \(n\),
Then, manipulations
\[
E(\tilde{L}_{n,\text{pos}}) \leq (\alpha_d + \delta/j)n^{1/d}.
\]
For any constant positive probability \( p_0 > 0 \), assume
\[
Pr(\tilde{L}_{n,\text{pos}} \geq (\alpha_d + \delta)n^{1/d}) \geq p_0
\]
Then,
\[
E(\tilde{L}_{n,\text{pos}})/n^{1/d} \geq (1 - \varepsilon_2 - p_0)(\alpha_d - \varepsilon_1) + p_0(\alpha_d + \delta)
\]
which for \( j \) big enough and \( \varepsilon_1, \varepsilon_2 \) small enough results in a contradiction.

To move from \( \tilde{L}_{n,\text{pos}} \) back to \( L_{n,\text{pos}} \), a process known as de-Poissonization, we "sandwich" \( L_{n,\text{pos}} \) between \( \tilde{L}_{n-n^{2/3},\text{pos}} \) and \( \tilde{L}_{n+n^{2/3},\text{pos}} \). We know that w.h.p. the number of points chosen in the Poisson process \( P(n - n^{2/3}) \) is less than \( n \). Therefore, if w.h.p. \( \tilde{L}_{n-n^{2/3},\text{pos}}/n^{1/d} \geq \alpha_d - \varepsilon \) then w.h.p. \( L_{n,\text{pos}}/n^{1/d} \geq \alpha - \varepsilon \). The same reasoning applies for the upper bound and we get that w.h.p. \( L_{n,\text{pos}}/n^{1/d} \to \alpha_d \).

We know that \( \alpha_d \leq e \), but we do not know yet that it is positive. A simple construction of Kingman, \([112]\) computes lower bounds for \( \alpha_d \).

Start with the origin and construct a chain using the following rule:

If \( A_i = (x_1, ..., x_d) \) is the last point added to the chain, then choose \( A_{i+1} \) to be the point which minimizes \( \sum_i x_i \) among all points \( A \) such that \( A_i \leq \text{pos} \ A \).

Let \( A_{i+1} = (x_1 + dx_1, ..., x_d + dx_d) \) and let \( dz = \sum_i dx_i/n^{1/d} \). The region \( \Delta_s \) given by \( dx_i \geq 0 \) and \( dz \leq s \) is a pyramid (simplex) with volume \( s^d/d! \). To see why, consider the region \( \Delta_1 \) given by \( x_i \geq 0 \) and \( \sum_i x_i = 1 \). Define \( y_i = \sum_{j \leq i} x_j \). Since the transformation from the \( x_i \) to the \( y_i \) is given by a linear transformation of determinant 1 it is volume preserving. The image consists of the set of vectors with coordinates \( y_i \) which satisfy \( 0 \leq y_1 \leq y_2 \leq ... \leq y_d \leq 1 \). The unit cube \( I^d \) is divided into \( d! \) equal volume subsets \( I^d_\pi \) indexed by permutations \( \pi \). The set \( I^d_\pi \) consists of all vectors \( y_i \) such that \( 0 \leq y_{\pi(1)} \leq ... \leq y_{\pi(d)} \leq 1 \). Putting these facts together we see that \( \text{vol} (\Delta_1) = \text{vol} (I^d/d!) = 1/d! \). Scaling by the constant \( s \) yields the desired conclusion since the image of \( \Delta_1 \) is \( \Delta_s \). We have that \( dz \geq s \) if none of the \( P(n) \) points of the Poisson sample will land in \( \Delta_s \). The probability for such an event is \( p_{ns^d/d!} = e^{-ns^d/d!} \). Differentiating \( Pr(dz \leq s) = 1 - e^{-ns^d/d!} \) we obtain the density function of \( dz \) which is \( f(s) = \frac{ns^{d-1}}{(d-1)!} e^{-ns^d/d!} \). We conclude that

\[
E(dz) = \int_0^\infty sf(s)ds = \int_0^\infty \frac{ns^d}{(d-1)!} e^{-ns^d/d!} ds
\]

Making the change of variable \( r = ns^d/d! \) we obtain after some simple manipulations

\[
E(dz) = n^{-1/d} (d!)^{1/d} \int_0^\infty r^{1/d} e^{-r} dr
\]
The increments \( dx_i \) are i.i.d. and their sum is \( dz \), hence
\[
E(dx_i) = \frac{1}{d} n^{-1/d}(dl)^{1/d} \int_0^\infty r^{1/d}e^{-r}dr
\]
Consider the point \( A_j = (x_1,j, \ldots, x_d,j) \). We have \( x_{i,j} = \sum_{l \leq j} dx_{i,l} \) where \( dx_{i,l} = x_{i,l} - x_{i,l-1} \) are all distributed like \( dx \). In addition, they are all independent since \( dx_{i,l+1} \) depends on the distribution of points in the region of points which satisfy \( A \geq pos A_l \), while all \( dx_{i,h}, h < l \), depend on the distribution of points in the disjoint region consisting of points \( (y_1, \ldots, y_d) \) such that \( \sum_i y_i < \sum_i x_{i,l} \). For any \( \varepsilon > 0 \), set
\[
J = J_{\varepsilon,n} = \frac{d}{d!^{1/d}}(\int_0^\infty r^{1/d}e^{-r}dr)^{-1}(1-\varepsilon)n^{1/d}
\]
By the law of large numbers, we conclude that w.h.p. for all \( i, x_{i,J} \leq 1 \), hence \( A_j \in I^d \), and we obtain a chain of length \( J \). We see that \( \alpha_d \geq J_{\varepsilon,n}/n^{1/d} = \frac{d}{d!^{1/d}}(\int_0^\infty r^{1/d}e^{-r}dr)^{-1}(1-\varepsilon) \) and hence \( \alpha_d \geq \frac{d}{d!^{1/d}}(\int_0^\infty r^{1/d}e^{-r}dr)^{-1} \). We note that as \( d \to \infty \) it is easy to check that \( \int_0^\infty r^{1/d}e^{-r}dr \to \int_0^\infty e^{-r}dr = 1 \) hence \( liminf_d \alpha_d \geq \frac{d}{d!^{1/d}} \to e \), the last inequality courtesy of Sterling’s formula. Since \( \alpha_d \leq e \) we have \( lim_d \alpha_d = e \).

To obtain the estimates on \( X_{n,pos} \), we follow the same proof plan but in a multiplicative rather than additive version. We start by showing that
\[
\beta_d = limsup_m \log(E(L_{m,pos}))/m^{1/d}
\]
is finite. We have \( E(X_{m,pos}) = \sum_{k=1}^m E(X_{m,k}) \leq m (Max_k E(X_{m,k})) \). Let \( \omega_d = Max_k (\frac{e}{k})^d b \). A simple computation of the derivative shows that the maximum is attained at \( b = 1 \) and we obtain \( \omega_d = e^d \). Using equation (42) and this calculation we see that for any \( \eta > 0 \) and \( m \) large enough \( (e^d - \eta)^{m^{1/d}} < Max_k E(X_{m,k}) < (e^d + \eta)^{m^{1/d}} \) from which we deduce that for \( m \) large enough \( E(X_{m,pos}) < (e^d + \eta)^{m^{1/d}} \) which implies that \( limsup_m \log(E(X_{m,pos}))/m^{1/d} \) is finite. The function log is concave on the positive numbers, which means that for any \( 0 \leq t \leq 1 \) and any \( a_1, a_2 > 0 \) we have \( t \log(a_1) + (1-t) \log(a_2) \leq \log(ta_1 + (1-t)a_2) \). This follows easily from the fact that the second derivative satisfies \( \log''(x) = \frac{1}{x^2} < 0 \).

Using the concavity of the logarithm function it is easy to prove inductively that for any finite probability distribution \( p_i \) and any set of values \( a_i > 0 \) we have \( \sum_i p_i \log(a_i) \leq \log(\sum_i p_i a_i) \). Applying this to the random variable \( X_{m,pos} \) we have \( E(\log(X_{m,pos})) \leq \log(E(X_{m,pos})) \), hence \( \beta_d = limsup_m E(\log(X_{m,pos}))/m^{1/d} \) is finite. The proof proceeds exactly as before. Previously we used the concatenation of chains of lengths \( L_{n_0,i} \) in the sub-cubes \( S_i \) to produce a chain of length \( \sum_i L_{n_0,i} \) in \( I^d \). In the new case if we have \( X_{n_0}(i) \) chains in sub-cube \( S_i \), then taking all concatenations will produce \( \prod_i X_{n_0,i} \) different chains in \( I^d \) and taking logarithms we get as before that \( \log(X_{n_0,i,pos}) \geq \sum_i \log(X_{n_0}(i)) \) and the proof goes through.
To see that $\beta_d$ is strictly positive, we note that by the first statement of the theorem, there exists w.h.p. a chain of length $(\alpha_d/2)n^{1/d}$. By taking all its sub-chains we see that w.h.p. $\log(X_{n,\text{pos}}) \geq \log(2)(\alpha_d/2)n^{1/d}$. q.e.d.

The values of $\alpha_d$, $d \geq 3$ are not known analytically, however $\alpha_2$ has been computed. At first, based on his numerical work, Ulam, [178], estimated that $\alpha = 1.7\ldots$ Later Baer and Brock, [36], using better numerics suggested that $\alpha = 2$. The determination of $\alpha_2$ required considerable effort, but in 1977, Vershik and Kerov, [179], and independently, Logan and Shepp, [123], were able to prove the following result, confirming the latter numerics.

**Theorem 23.** If $d = 2$, then, w.h.p. $|L_n - 2\sqrt{n}| < \varepsilon \sqrt{n}$.

We want a version of theorem 22 for Minkowski space.

**Theorem 24.** Let $I^d_M = I(A, B)$ be an interval in $d$ dimensional Minkowski space, where $A <_M B$. Let $L^d_n$, respectively $\tilde{L}^d_n$, denote the length of the longest chain w.r.t. $\leq_M$ among $n$, respectively $P(n)$, uniformly sampled points in $I^d_M$. Let $X^M_n$, $\tilde{X}^M_n$ denote the total number of chains among $n$, $P(n)$, uniformly sampled points respectively. Then,

1) There is a constant $\alpha^M_d > 0$ such that for any $\varepsilon > 0$, w.h.p.
$$|L^d_n - \alpha^M_d n^{1/d}| < \varepsilon n^{1/d}$$

The exact same estimate holds for $\tilde{L}^d_n$.

2) There is a constant $\beta^M_d > 0$ such that for any $\varepsilon > 0$, w.h.p.
$$(\beta^M_d - \varepsilon)^{n^{1/d}} < X^M_n < (\beta^M_d + \varepsilon)^{n^{1/d}}$$

The exact same estimate holds for $\tilde{X}^M_n$.

**Proof:** The proof follows the same outline as the proof of theorem 22. The basic idea is that we can find Minkowski matrices $M$ and $\tilde{M}$ so that $C \leq_M D$ implies $C \leq_{\text{pos}} D$ which in turn implies $C \leq_M D$ for any $C, D$. Let $A = (0, 0, \ldots, 0)$ and $B = (1, 1, \ldots, 1)$. We recall that for the Minkowski matrices $D_r$, the causal cone was given by connecting the origin to points in a ball of radius $r$ in the hyperplane $H$ given by the equation $t = 1$ and centered at $(1, 0, \ldots, 0)$. By rotating and scaling Minkowski space we can find a Minkowski metric $M$ for which the cone of causal vector will be obtained by joining the origin to points of radius $\hat{r}$ in a ball in the hyperplane $\tilde{H}$, given by $\sum_{i=1}^d x_i = d/2$ and centered at $(1/2, \ldots, 1/2)$ the midpoint between $A$ and $B$. Obviously $A <_M B$. Finally, we assume that the ball is of a small enough radius so that it is entirely contained in the unit cube. A ball of radius $1/2$ will do. We observe that $0 \leq_{\text{pos}} D$ iff $D$ is on the ray connecting the origin with a point on $\tilde{H} \cap I^d$. Since we have chosen the ball to be contained in this intersection we see that $0 \leq_M D$ implies $0 \leq_{\text{pos}} D$, and by translation this holds for all $C, D$. Taking the radius of the ball to be large...
enough so that it contains \( \tilde{H} \cap I^d \) yields \( \tilde{M} \). It is easy to check that the points in \( \tilde{H} \cap I^d \) furthest away from \((1/2, \ldots, 1/2)\) have a single non-zero coordinate whose value is \( d/2 \). A simple calculation then shows that taking the radius to be \( \frac{d+1}{2} \) is sufficient. From our choices we conclude that if \( D \leq_M E \) then \( D \leq_{pos} E \) and that consequently we also have \( I^d_M \subset I^d \). Choosing uniformly \( P(n) \) points in \( I^d_M \) can be considered as a restriction to \( I^d_M \) of the process of choosing uniformly \( P(\tilde{n}) \) points where \( \tilde{n} = \frac{\text{vol}(I^d)}{\text{vol}(I^n)} \) points in \( I^d \). Since any chain in Minkowski space is also a chain in \( I^d \) w.r.t. \( \leq_{pos} \), we have, \( E(\tilde{L}^M_n) \leq E(\tilde{L}_{\tilde{n},pos}) \). From this, using the proof of theorem 22, we deduce the finiteness of \( \text{limsup}_m E(L^M_n)/m^{1/d} \) and \( \text{limsup}_m E(\log(X^M_n))/m^{1/d} \).

The role of the sub-cubes \( S_j \) is played by the intervals \( I_M(A_j, A_{j+1}) \) for \( 0 \leq j \leq l - 1 \), where \( A_j = (j/l, \ldots, j/l) \). We have the same volume ratios \( \frac{\text{vol}(I_M(A, B))}{\text{vol}(I^n_M(A_j, A_{j+1}))} = l^d \) and the partial order is scale invariant. The rest of the proof follows the exact same steps as before.

The positivity of \( \alpha^M_d \) can be shown by considering \( \tilde{M} \). Since \( A \leq_M B \) implies \( A \leq_{pos} B \) we get in particular \( I^d \subset I^n_M(A, B) \). Sampling \( P(\tilde{n}) \) points in \( I^d_M(A, B) \), will produce a \( P(\frac{\text{vol}(I^d)}{\text{vol}(I^n_M(A, B))} n) \) sampling of \( I^d \) and we can use the positivity result about the latter to see that \( \alpha^M_d = \alpha^M_d \) is positive. q.e.d.

It is possible to prove the theorem directly without reduction to the \( \leq_{pos} \) case. This was done for a family of partial orders which includes Minkowski space in [36]. One needs to replace the simple estimate \( \frac{1}{k^{1/d}} \) for the probability that a set of size \( k \) in \( \leq_{pos} \) will be a chain, with an appropriate upper bound estimate for the same probability in the case of \( \leq_M \). This is done inductively by conditioning on the location of the middle point of the chain, say \( E \), and considering inductively the probability estimates for the points smaller than \( E \) and those larger than \( E \).

The analogue of Kingman’s greedy construction also needs to be developed. The next point in the chain is taken to be the point \( E \) that maximizes the remaining volume \( I^n_M(E, B) \).

Using these methods, Bollobas and Brightwell showed that

\[
\int_0^\infty t^{1/d}e^{-t}dt \leq \alpha^M_d \leq 2^{1-1/d} e(\ell d)^{1/d} \frac{d}{d}
\]

In particular as \( d \to \infty \) then \( \alpha^M_d \to 2 \).

**Proof of theorem 19**: We first show the lower bound that for any \( \varepsilon > 0 \), w.h.p., \( L_n > (\alpha_d \omega_d \text{diam}(D) - \varepsilon)n^{1/d} \). Given some \( \varepsilon_1 > 0 \), consider a causal curve \( \gamma : [a, b] \to D \) such that \( \ell(\gamma) = \text{diam}(D) - \varepsilon_1 \) and which is contained in the interior of \( D \). By our assumptions on \( D \), such a curve exists. We fix a finite covering of \( \gamma \) by normal open sets \( U_i \subset D \), \( 1 \leq i \leq m \), and points \( A_i = \gamma(a_i) \) with \( a = a_0 < a_1 < \ldots < a_m = b \) such that \( \gamma_i = \gamma([a_{i-1}, a_i]) \subset U_i \).
Let $d_i$ denote the usual Euclidean distance on $U_i$ and let $\ell_{euc,i}$ denote the associated length function, which by the Lifschitz property is finite on causal curves.

For $s \in [a, b]$, define $i(s)$ to be the index such that $a_{i(s) - 1} \leq s < a_{i(s)}$. For any given $x = \gamma(s)$, $s \in [a, b]$ and any $\delta > 0$, we can find, by lemma 10, a neighborhood $N_{s, \delta} \subset U_{i(s)}$ of $x$, which is $\delta$-normal with respect to the Euclidean metric $d_i(s)$.

Given $\delta > 0$ we refine the set of points $A_i$ to a finite set of points $B_j = B_{j, \delta}$ by enlarging the set $a_i$ to a finite set $s_i$ such that $\gamma(s_j) = B_j$ and the open sets $N_i = N_{s_j, \delta}$ form a cover of $\gamma$. This is possible by the compactness of $\gamma$ and the existence of $\delta$-normal neighborhoods. We let $i(j) = i(s_j)$. We further enlarge the set $s_j$ to a finite set $t_k$ such that each successive pair of points among $C_k = \gamma(t_k)$ belongs to some neighborhood $N_{j(k)}$.

We recall that being a normal neighborhood means that there exists two Minkowski metrics $g_{maj,j}$ and $g_{min,j}$ on $N_j$ such that for any vector $v$ centered at $x \in N_j$ we have $vg_{min,j}v^t < vg_{maj,j}v^t$. We also have the condition that the set $N_j$ is causally convex w.r.t. $g_{maj,j}$ and with respect to $g$ using strong causality. The $\delta$-normality condition states that for any pair of points $E, F \in N_j$, either $1 - \delta < d_{min,j}(E, F)/d_{maj,j}(E, F)$ or $d_{maj,j}(E, F) < \delta d_{j}(E, F)$. Given an index $k$ we use the abbreviations $j = j(k)$ and $i = i(j(k))$. For a given $\varepsilon_1$ we let $K = K_{\varepsilon_1}$ be the set of indices $k$ such that $d_{maj,j}(C_k, C_{k+1}) \geq \varepsilon_1 d_i(C_k, C_{k+1})$. For each $k \in K$ we consider the interval $I_{min,j}(B_k, B_{k+1})$. Let $\text{vol}_{min,k} = \text{vol}_{min,j}(I_{min,j}(C_k, C_{k+1}))$ and likewise $\text{vol}_{maj,k} = \text{vol}_{maj,j}(I_{maj,j}(C_k, C_{k+1}))$. The number of sampled points from $I_{min,j}(C_k, C_{k+1})$ will be $P(\text{vol}_{min,k,n})$ and by theorem 24 for any $\varepsilon_2 > 0$ and all $k \in K$, w.h.p. the longest chain in this set w.r.t. $\leq_{min,i}$ is of size at least

$$(\alpha_d - \varepsilon_2)(\text{vol}_{min,k,n})^{1/d} = (\alpha_d - \varepsilon_2)\omega_d d_{min,j}(C_k, C_{k+1})n^{1/d}$$

$$\geq (1 - \delta)(\alpha_d - \varepsilon_2)d_{maj,k}(C_k, C_{k+1})n^{1/d}$$

Concatenating the longest chains for the different values of $k \in K$ and noting that a chain w.r.t. $\leq_{min,j}$ is also a chain w.r.t. $\leq_S$ we obtain a chain in $D$ of size at least

$$(1 - \varepsilon_1)(\alpha_d - \varepsilon_2)\omega_d(\sum_{k \in K} d_{maj,k}(C_k, C_{k+1}))n^{1/d}$$

We have

$$\sum_{k \notin K} d_{maj,j}(C_k, C_{k+1}) \leq \varepsilon_1(\sum_{k \notin K} d_i(C_k, C_{k+1}) \leq \varepsilon_1(\sum_{i} \ell_{euc,i}(\gamma_i))$$
where \( \gamma_i \) is the image under \( \gamma \) of the interval \([a_{i-1}, a_i]\). Note that this bound is independent of \( \delta \). We conclude that
\[
\sum_{k \in K} d_{\text{maj},j}(C_k, C_{k+1}) = \sum_{k} d_{\text{maj},j}(C_k, C_{k+1}) - \sum_{k \notin K} d_{\text{maj},j}(C_k, C_{k+1}) \geq \ell(\gamma) - \varepsilon_1(\sum_i \ell_{\text{eucl},i}(\gamma_i))
\]
where we have used the definition of \( \ell \) constructed above. Taking \( \varepsilon_1, \varepsilon_2 \) and \( \delta \) small enough we conclude that w.h.p. there is a chain of size at least
\[(\alpha d \omega d \ell(\gamma) - \varepsilon)n^{1/d}\]
To prove the upper bound, choose the \( C_k \) as above so that
\[
\sum_k T_{\text{maj},j}(C_k, C_{k+1}) \leq \ell(\gamma) + \varepsilon/2
\]
Such a set exists by the definition of the length function. As in the proof of theorem 15, for each \( k \) choose points \( D_k, E_k \in N_j \) such that \( I_{\text{max},j}(D_k, E_{k+1}) \) contains \( I_{\text{max}}(C_k, C_{k+1}) \) as an open set and such that \( d_{\text{max},j}(D_k, E_{k+1}) < (d_{\text{max},j}(C_k, C_{k+1}))/\varepsilon/4|K| \), where \(|K|\) denotes the size of the set \( K \). Consider the union \( W_{\gamma,\varepsilon} = \bigcup_k I_{\text{max},j}(D_k, E_{k+1}) \). This is an open set which contains \( \gamma \). Any chain \( G \) in \( W_{\gamma} \) is the union of its restrictions \( G_k \) to \( I_{\text{max},j}(D_k, E_{k+1}) \), which by strong causality are themselves chains w.r.t. \( \leq_{\text{max},j} \). By theorem 24, for any \( \delta > 0 \), the longest chain w.r.t. \( \leq_{\text{max},j} \) in \( I_{\text{max},j}(D_k, E_{k+1}) \) has length at most \((\alpha d \omega d + \delta / 4) d_{\text{max},j}(D_k, E_{k+1}) n^{1/d}\) and hence, after summing, w.h.p. the longest chain in \( W_{\gamma,\varepsilon} \) has length at most
\[(\alpha d \omega d + \delta)(\ell(\gamma) + 3\varepsilon/4)n^{1/d} < (\alpha d \omega d \ell(\gamma) + \varepsilon)n^{1/d} \leq (\alpha d \omega d \text{diam}(D) + \varepsilon)n^{1/d}\]
the next to last inequality for \( \delta \) small enough.

Any chain among the sampled points is by definition part of a causal curve. Since \( C_D(D, D) \), the set of causal curves in \( D \) is compact in the \( C_0 \) topology there is a finite set of causal curves \( \gamma_i, i = 1, \ldots, m \) such that \( W_{\gamma_i,\varepsilon} \) cover \( C_D(D, D) \) and hence inequality (43) holds w.h.p. for all chains among the sampled points.

Let \( V_{\gamma,\varepsilon} \) be the intersection of \( W_{\gamma,\varepsilon} \) with the open set of points \( x \in D \) such that \( d_S(x, \gamma) < \varepsilon \). Consider the union of all sets \( V_{\gamma,\varepsilon} \), where \( \gamma \) is a maximal curve. This is an open set in \( C_D(D, D) \) so the complement \( A \) is a closed set in \( C_D(D, D) \) and hence compact. By definition there are no maximal curves in \( A \) and hence the maximal length of a curve in \( A \) which exists by compactness is strictly less than \( \text{diam}(D) \), say \( \text{diam}(D) - \delta \). Covering \( A \) by a finite number of open sets of the form \( V_{\gamma_i,\delta/2} \), we see that w.h.p. the longest chain which is not in an \( \varepsilon \) neighborhood has size at most \((\alpha d \omega d - (\delta / 4)) \text{diam}(D)n^{1/d}\), hence, by the high probability lower bound on the size of a maximal chain it will not be maximal.

The proofs for the number of chains are precisely the same. \( q.e.d. \)
8. Modeling disk scheduling and airplane boarding via Lorentzian geometry

8.1. Disk scheduling and Lorentzian geometry. Following the previous section, we can state a result that summarizes our analysis of disk scheduling with a linear seek function in the language of Lorentzian geometry.

**Theorem 25.** Consider a disk drive with seek time function \( f(t) = t \). Let \( D \) be the domain \( 0 \leq r, \theta \leq 1 \) in the cylinder \( C \), representing disk locations. Let \( p(\theta, r) \) be a request probability distribution on \( D \). Let \( g \) be the Lorentzian metric induced by \( 2p(\theta, r) \). Let \( \text{diam}(D) \) denote the maximal length of a causal curve in \( D \). Let \( \bar{R} \) be a set of \( n \) requests sampled i.i.d. from the density \( p \) and let \( ST(\bar{R}) \) be the number of disk rotations needed to service all the requests in \( \bar{R} \) with an optimal policy. Then:

1) Up to a bounded additive constant, the ABZ algorithm is optimal.

2) For all \( \varepsilon > 0 \), w.h.p.

\[
\left| \frac{ST(\bar{R}) - \text{diam}(D)}{\sqrt{n}} \right| < \varepsilon \sqrt{n}
\]

3) When \( p(\theta, r) = p(r) \) depends only on \( r \) we have

\[
diam(D) = \sqrt{2} \int_0^1 \sqrt{p(r)} dr
\]

**proof:** We have proved the first two items, the second being a direct consequence of the analysis of the ABZ algorithm and theorem 19. For the third item, we need to show that the maximal curves of the model are the lines where \( t \) is fixed. It is easy to verify that the length of such a curve is given by the expression in (45). Consider any curve \( \gamma \) which contends to be a maximal curve. let \( E_1, \ldots, E_k \), with \( E_i = (t_i, r_i) \), be a set of points on \( \gamma \). The corresponding Riemann sum length estimate for \( C \) will be \( \sum_i \sqrt{2p(r_i)}(dr_i^2 - dt_i^2) \leq \sum_i \sqrt{2p(r_i)}dr_i \). The right hand side is the length estimate associated with the points \( E_i' = (t_0, r_i) \) on the curve given by fixing the value of \( t \) to be \( t_0 \). We conclude that the latter curve is at least as long. \( \text{q.e.d.} \)

8.2. Airplane boarding and Lorentzian geometry. Consider the stochastic airplane boarding process with parameters \( n, h, l, w, p \), as defined previously in section 5.

**Definition 69.** We define the **blocking potential** \( \alpha(q, r) \) by the formula

\[
\alpha(q, r) = \int_r^1 p(q, z) dz
\]
To understand the meaning of \( \alpha(q,r) \), we consider a passenger \( A \) whose coordinate representation is \((q,r)\). The blocking potential measures the density of passengers which are closely behind passenger \( A \) in the queue, and whose assigned row is beyond row \( r \). Stated otherwise, \( \alpha(q,r) \) indicates how many passengers are standing behind passenger \( A \), waiting to pass by him/her on the way to their assigned seats. Such passengers play an important role in the precedence relation, since they cause the backlog which allows passenger \( A \) to block other passengers. Note that among the input parameters, \( \alpha \) depends only on the boarding policy function \( p \).

**Definition 70.** The congestion parameter \( k \), is defined as

\[
k = \frac{hw}{l}.
\]

To understand the meaning of \( k \), consider all the passengers in the airplane. If we were to let all passengers queue in a long line, then according to our parameters, the total length of the line would be \( nw \). On the other hand the length of the airplane aisle is \( sl = nl/h \), where \( s \) is the number of rows and \( l \) the distance between successive rows. The congestion parameter \( k \) is the ratio of the two numbers. If \( k < 1 \) then all passengers can stand at the same time in the airplane aisle. In the more realistic situation where \( k > 1 \) then only \( 1/k \) of all passengers can stand in the aisle at the same time. We can estimate \( k \) by observing passengers when they are waiting to leave the airplane. The aisle is usually completely occupied, but most passengers are still seated since there is no room for them in the aisle.

If the distance between rows is enlarged or the number of passengers per row is reduced, then \( k \) will decrease. In this sense, \( k \) is a measure of congestion. Note that \( k \) depends only on the airplane related input parameters and not on the airline policy.

**Definition 71.** We assign to the boarding process with parameters \( h,l,w,p \) the space-time domain \( S_{k,p} \) consisting of the boarding metric on the \((q,r)\) unit square

\[
4p(q,r) \begin{pmatrix} k\alpha(q,r) & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}
\]

We denote by \( L = L_{n,p,k} \) the length of the longest chain of a discrete space-time associated with \( n \) points sampled from \( S_{k,p} \). We denote by \( \tilde{L} = \tilde{n}_{n,p,k} \) the longest chain among \( P(n) \) sampled points.

It is easy to verify that \( S_{k,p} \) is strongly causal since the \( q \) coordinate is non-decreasing along any chain and for elements in the chain with the same \( q \) coordinate, \( r \) is non-decreasing.

As in the case of disk scheduling, there are two properties of the boarding metric which relate it to the boarding process:

(M1) The volume form of the metric which is \( 2p(q,r) dq dr \) is proportional to the passenger density distribution \( p(q,r) dq dr \).
9. Computing the maximal curve

(M2) The blocking partial order among passengers during the boarding process almost coincides with the causal partial order relation.

To establish, heuristically, the second property, consider passengers represented by $X = (q,r)$ and $X' = (q + dq, r + dr)$, $dq > 0$. Consider the time when passenger $X$ arrives at his/her designated row. All passengers with row numbers beyond $r$, which are behind passenger $X$ in the queue but in front of passenger $X'$, will occupy aisle space behind passenger $X$. The number of such passengers is roughly $\alpha dq n$. Each such passenger occupies $w/l$ units of aisle length, where we take the basic aisle length unit to be the distance between successive rows. The row difference between $X$ and $X'$ is $-(1/h)dr n$. We conclude that passenger $X$ is blocking passenger $X'$, via the passengers between them, roughly when $k\alpha dq \geq -dr$, a condition which coincides (together with $dq > 0$) with the causal relation induced by the metric $dq(k\alpha dq + dr) \geq 0$.

As a result of these properties we have the following basic modeling result whose rather tedious proof (formally justifying the heuristic argument for M2) is given in the appendix.

**Theorem 26.** Let $L_{n,h,l,w,p}$ be the boarding time random variable and $L_{n,k,p}$ be the longest chain among $n$ points sampled from $S_{k,p}$. Then, for all $\varepsilon > 0$, w.h.p., we have

$$|L_{n,h,l,w,p} - L_{n,k,p}| \leq \varepsilon \sqrt{n}$$

9. Computing the maximal curve

We would like an equation for the maximal curve so we can compute its length and find the diameter of $D$. For simplicity we will consider only the $d = 2$ dimensional case and also restrict ourselves to a local calculation in the neighborhood of a point. The generalization to higher dimensions is easy.

As we have shown before, locally in a neighborhood of a point, causal curves in a $d$ dimensional space-time can be considered as Lifschitz functions from a parameter which we will call $x$ (rather than $t$ that we used before) into $d - 1$ dimensional space, which for $d = 2$ is $\mathbb{R}$. Let the curve be given as $y = y(x)$. The length of the curve will be given by a formula of the form

$$\int_{a}^{b} \sqrt{M(x,y) v^t} dx$$

where $v = (1, y')$ is the tangent vector to the curve.

**Definition 72.** Consider differentiable curves $y = y(x)$ which start at some fixed point $(a, y(a))$ and end at another fixed point $(b, y(b))$. An expression of the form

$$\int_{a}^{b} L(x, y, y') dx$$

where $L$ is some function which depends smoothly on the variables is called a **functional** or an **action**.
The calculus of variations provides a necessary condition for a curve to maximize or minimize a functional. The condition is analogous to the vanishing derivative criterion for an extremal value of a function on the reals.

Sometimes the function $L$ may be defined only for a restricted family of differentiable curves. Such is the case for length which is only defined for causal curves. We say that $F$ is unrestricted at a point $x, y, y'$ if it is still defined if we change these values by a small amount, say less than some $\varepsilon > 0$.

For causal curves and the length functional this means that the tangent to the curve is time-like and not only causal. It also means that $a < x < b$ and that $(x, y)$ can be connected by time-like curves to $(a, y(a))$ and to $(b, y(b))$. Another form of a constraint is that the curve has to stay within a domain $D$. If the point $(x, y)$ is on the boundary of the domain the functional will be constrained at that point.

The following result can be viewed as the basic result of the calculus of variations

**Theorem 27.** Assume that the curve $y_0(x)$ maximizes or minimizes the functional $G(y) = \int_a^b L(x, y, y')dx$ among curves passing between $(a, y_0(a))$ and $(b, y_0(b))$ and assume that $y_0$ is unconstrained at $(x, y_0(x))$, then, $y_0(x)$ must satisfy the differential equation

\[
\frac{\partial L}{\partial y} - \frac{d}{dx} \left( \frac{\partial L}{\partial y'} \right) = 0
\]

This equation is known as the Euler-Lagrange equation

**Proof:** Let $y_0 = y_0(x)$ be a maximizer or minimizer of the functional $F$. Let $z(x)$ be some continuously differentiable function of $x$ with $z(a) = z(b) = 0$. Consider the function $y_\delta(x) = y(x) + \delta z(x)$. For small enough $\delta$ we will have $|\delta z|, |(\delta z)'| < \varepsilon$ and so by our assumption, the functional $F$ is defined on $y_\delta$. Consider the 1-variable function $G(\delta) = \int_a^b L(x, y_\delta, y_\delta')dx$. By our assumption $G(0) = G(y_0)$ is a minimum or maximum of $G$, hence $G'(0) = 0$. Computing we have $\frac{dG}{d\delta} = \frac{\partial G}{\partial x} \frac{dx}{d\delta} + \frac{\partial G}{\partial y} \frac{dy}{d\delta} + \frac{\partial G}{\partial y'} \frac{dy'}{d\delta}$. We have $\frac{dx}{d\delta} = 0$, $\frac{dy}{d\delta} = z$ and $\frac{dy'}{d\delta} = z'$. Plugging these equations and exchanging the integral with the derivative since the integrand is smooth we get $\frac{dG}{d\delta} = \int_a^b \left( \frac{\partial L}{\partial y} z + \frac{\partial L}{\partial y'} z' \right) dx$. Integrating the second integrand by parts and recalling that $z(a) = z(b) = 0$ we get $\int_a^b \frac{\partial L}{\partial y'} z' dx = - \int_a^b \frac{d}{dx} \left( \frac{\partial L}{\partial y} \right) z dx$ and consequently $0 = \frac{dG}{d\delta}(0) = \int_a^b \frac{\partial L}{\partial y} (y_0) - \frac{d}{dx} \left( \frac{\partial L}{\partial y} \right) (y_0) \frac{dy}{d\delta} z dx$ and this equality is true for any continuously differentiable $z$. If $\frac{\partial L}{\partial y} (y_0) - \frac{d}{dx} \left( \frac{\partial L}{\partial y} \right) (y_0) \neq 0$ for some point $x_0$ then by continuity it will not vanish in a neighborhood of $x_0$ and we can choose some function $z$ which is positive on the neighborhood and vanishes outside it that will produce a contradiction. $q.e.d.$
In some cases, the function $L$ does not depend on one of the variables explicitly. In such cases we obtain a conservation law.

**Theorem 28.** Assume that the curve $y_0(x)$ maximizes or minimizes the functional $G(y) = \int_a^b L(y, y') dx$ among curves passing between $(a, y_0(a))$ and $(b, y_0(b))$ and assume that $y$ is an unconstrained solution for the functional. Let $H(x) = y' \frac{\partial L}{\partial y'} - L$. Then, $y_0(x)$ must satisfy the differential equation

$$H(x) = c$$

for some constant $c$. This equation is known as the Beltrami equation. The constant $c$ is known as a constant of motion or a conserved quantity. The quantity $H$ is called the Hamiltonian or energy.

**Proof:** By the Euler-Lagrange equation we have $\frac{d}{dx} \left( \frac{\partial L}{\partial y'}(y_0) \right) = \frac{\partial L}{\partial y}(y_0)$, hence $(\frac{\partial L}{\partial y'}(y_0))' = \frac{\partial L}{\partial y}(y_0)$. Consequently

$$H(y_0)' = y_0' \frac{\partial L}{\partial y}(y_0) + y_0'' \frac{\partial L}{\partial y'}(y_0) - \frac{\partial L}{\partial y'}(y_0)y_0' - \frac{\partial L}{\partial y}(y_0)y_0'' = 0$$

as required. q.e.d.

**9.1. Existence of solutions.** We obtained above a differential equation, the Euler-Lagrange equation, that we wish to apply to find maximal curves. In general, causal curves are not always differentiable (although they are differentiable almost everywhere). However, in the cases of interest to us it can be shown using the existence and uniqueness theorem for ordinary differential equations, the Picard-Lindelof theorem, that in a small enough neighborhood $N_x$ of a point $x$ in a strongly causal space-time $S$, if $x < S y$, there is a differentiable maximal curve starting at $x$ and ending at $y$.

**9.2. Computing the boarding time of the random boarding policy.** We apply the above equations to find the length of the maximal curve in the Lorentzian domain corresponding to the random boarding policy $F$. Recall from section 5 that for the random boarding policy the joint row/queue distribution $p$ is uniform, namely, $p(q,r) = 1$. Since $p$ is uniform, we have by formula (46) $\alpha(q,r) = 1 - r$. We parameterize a causal curve, $\phi$, via the $q$-coordinate and write the curve accordingly in the form $\phi = \phi(q)$. The functional $L(\phi)$ takes the form

$$L(\phi) = \int_0^1 \sqrt{\phi' + k(1-\phi)} dq.$$  \hspace{1cm} (51)

Since the functional $L(\phi)$ does not depend explicitly on $q$, the Euler-Lagrange equation associated with the functional degenerates to the Beltrami equation

$$\phi' \frac{\partial L}{\partial \phi'} - L = c,$$  \hspace{1cm} (52)
or, explicitly

\[ \frac{\phi'}{2\sqrt{\phi' + k(1 - \phi)}} - \sqrt{\phi' + k(1 - \phi)} = c \]

The general solution of the equation is

\[ \phi(q) = c_1 e^{2kq} + c_2 e^{kq} + 1. \]

Given such a solution in a range \( a \leq q \leq b \), the value of the length of \( \phi \) is

\[ L(\phi) = (e^{kb} - e^{ka}) \sqrt{\frac{c_1}{k}}. \]

Since in this case the unit square is the interval between the points \((0, 0)\) and \((1, 1)\) a maximal curve \( \phi(q) \) will satisfy \( \phi(0) = 0, \phi(1) = 1. \) Assuming the maximal curve does not pass on the boundary of the unit square then placing these boundary conditions, we obtain the solution

\[ \phi(q) = \frac{e^{2kq}}{e^k - 1} - \frac{e^k}{e^k - 1} e^{kq} + 1. \]

The solution remains within the unit square for \( 0 \leq q \leq 1 \) when \( k \leq \ln(2) \). Applying the functional to the solution in this case, we obtain the diameter

\[ B(F, k) = \sqrt{\frac{e^k - 1}{k}} \]

for \( k \leq \ln(2) \).

When \( k > \ln 2 \), the solution (56) is not contained in the unit square anymore and we have to consider the boundary of the square, where the curve is constrained and so the Euler-Lagrange equation does not hold. We claim that the curve, which maximizes \( L \) subject to the boundary conditions \( \phi(0) = 0, \phi(q_0) = 0, 0 < q_0 \leq 1 \), is given by the \( q \)-axis segment \( \phi(q) = 0 \) between the endpoints. To see this, we can compare the functional \( L \) to a simpler functional \( \tilde{L}(\phi) = \int_0^1 \sqrt{\phi'^2 + k\phi'^2} dq \). Obviously, \( \tilde{L}(\phi) \geq L(\phi) \) for any curve \( \phi(q) \) in the unit square. It is easy to verify that the Euler-Lagrange equation associated with \( \tilde{L} \) is \( \phi'' = 0 \), and so any line segment is a solution. The claim follows by noting that, for \( \phi(q) = 0 \), the values of \( L \) and \( \tilde{L} \) coincide.

As we noted above a maximal curve has to be differentiable. It can be shown, via approximations using Minkowski metrics (where maximal curves are given by straight lines) that this will also be true on the boundary in our case. We conclude that the maximal curve \( \phi(q) \) consists of a segment \( \phi_1 \) of the form \([0, q_0]\) on the \( q \)-axis, followed by a curve \( \phi_2 \) which solves the Euler-Lagrange equation between the points \((q_0, 0)\) and \((1, 1)\). By the differentiability of \( \phi \) we need \( \phi_2'(q_0) = 0 \). We solve the Euler-Lagrange equation with boundary conditions \( \phi_2(q_0) = 0 \) and \( \phi(1) = 1 \), requiring in addition that \( \phi_2'(q_0) = 0 \), which leads to \( q_0 = \frac{k - \ln(2)}{k} \). Plugging the resulting
maximal curve $\phi$ into the functional leads to a diameter of

$$B(F, k) = \sqrt{k} + \frac{1 - \ln 2}{\sqrt{k}}, \quad k > \ln(2).$$

The maximal curves are plotted for various values of $k$ in Figure 1.

10. Analysis of back-to-front boarding policies

Our aim in this subsection is to prove a lower bound on the boarding time of any back-to-front policy, when the congestion parameter satisfies $k \geq 1$. We note that in practice this is the case.

Let $B(\bar{r}, k)$ be the diameter of the space-time associated with the back-to-front boarding policy $F_{\bar{r}}$ defined in 34 and congestion parameter $k$.

Our strategy for obtaining lower bounds on $B$ is to present a particular causal curve $\varphi$ with a large value $L(\varphi)$.

Let $b(q)$ denote the piecewise linear function defined by the union of the bottom edges of squares $S_i$ which are defined in the paragraph following definition 34. The curve $b(q)$ is given by $b(q) = r_i$ for $1 - r_{i-1} \leq q < 1 - r_i$, $i = 1, 2, \ldots, m$.

We have $\alpha(q, b(q)) = 1$, since $p(q, r) = 0$ for $r < b(q)$.

Given a partition $\bar{r}$, given by $1 = r_0, r_1, \ldots, r_{m-1}, r_m = 0$, and an index $j \in \{1, 2, \ldots, m\}$, we define a piecewise linear continuous function $\varphi_{(\bar{r}, j)}(q) = \varphi_j(q)$ as follows. The variable $q$ is in the range $[0, 1 - r_j]$. The graph of the function $\varphi_j$ is composed of $h = h(j, \bar{r})$ linear segments, $\psi_1, \ldots, \psi_h$ where $h$ is an integer, $1 \leq h \leq 2 \cdot j$.

The segments are of two types. A segment $\psi$ of the first type is a horizontal segment (that is, a segment with slope 0), and it is necessarily a subsegment of some bottom edge $B_i$ for an index $i$ between 1 and $j$. Moreover, the segment $\psi$ contains the left endpoint $(1 - r_{i-1}, r_i)$ of the segment $B_i$. Finally, the right-most segment $\psi_h$ is of the first type and consists of the entire bottom edge $B_j$ of the square $S_j$.

A segment $\psi$ of the second type is a segment with slope $(-k)$ that ends in a point $(1 - r_{i-1}, r_i)$ for some index $i$, $1 \leq i \leq j$. Moreover, for all values of $q$ for which $\psi(q)$ is defined, the inequality $\psi(q) \leq b(q)$ holds. Fix an index $j$, $1 \leq j \leq m$. The curve $\varphi_j(q)$ is the unique piecewise linear continuous curve in which segments of the first and second types alternate. The sequence $(\psi_1, \psi_2, \ldots, \psi_h)$ is called the segment decomposition of the curve $\varphi_j$, see figures 2 and 3 for an illustration.

Consider an index $j'$, $j < j' \leq m$. If $(\bar{q}, \bar{r})$ is a point of $\varphi_{j'}$, which belongs to the bottom edge $B_j$ of the square $S_j$ then the curves $\varphi_j$ and $\varphi_{j'}$ coincide in the range $[0, \bar{q}]$. The following theorem provides a lower bound on the diameter.

**Theorem 29.** If $k > 1$ then for any partition $\bar{r}$, we have

$$L(\varphi_{(\bar{r}, m)}) \geq \sqrt{k - 1}$$

In particular $B(\bar{r}, k) \geq \sqrt{k - 1}$. 
Figure 2. The curve $\varphi_m$ with $m = 5$ and $h = 6$. The segments of the second type are depicted by diagonal lines, and the segments of the first type are depicted by horizontal lines. These segments alternate.

Proof: For a fixed partition $\bar{r}$, let $\Omega = \{\varphi_1, \varphi_2, \ldots, \varphi_m\}$ be the family of $m$ curves as above. For curves $\varphi \in \Omega$ there is a combinatorial description of the functional $L(\varphi)$. Specifically, consider a curve $\varphi \in \Omega$, and let $\psi_1, \psi_2, \ldots, \psi_h$, be its decomposition into linear segments. By definition, $L(\varphi) = \sum_\ell L(\psi_\ell)$. Since $p(q, r) = 0$ for all points $(q, r)$ with $0 \leq q \leq 1$ and $r < b(q)$, it follows that for any segment $\psi$ of the second type, we have $L(\psi) = 0$ and consequently $L(\varphi_j)$ is the sum of the contributions of the segments of the first type. Consider now a segment $\psi$ of the first type defined on an interval $[q_1(\psi), q_2(\psi)]$, with $q_1 = q_1(\psi) = 1 - r_{i-1}$. The segment $\psi$ is contained in the bottom edge $B_i$ of the square $S_i$, for some index $i$ between 1 and $m$, and so $q_2 \leq 1 - r_i$.

since $\alpha = 1$ for all points in $\psi$ and $p = \frac{1}{r_{i-1} - r_i}$ we obtain

\begin{equation}
L(\psi) = \sqrt{k(q_2 - q_1)} \sqrt{\frac{1}{r_{i-1} - r_i}}
\end{equation}

We will prove the theorem by induction on $m$. Consider the case $m = 1$. In this case there is a unique partition $\bar{r} = (r_0 = 1, r_1 = 0)$ which corresponds to random boarding. For this partition, $S_1$ is the entire unit square, and
the density function \( p(q, r) \) is given by \( p(q, r) = 1 \) for all points \((q, r)\) in the square \( S_1 \).

In this case, by definition \( \varphi_1(q) = 0 \) for all \( q, 0 \leq q \leq 1 \), and by equation (59), \( L(\varphi_1) = \sqrt{k} \).

Assume the theorem has been proved for all partitions with \( m \) groups or less.

Let \( \bar{r} = (r_0 = 1, r_1, \ldots, r_m, r_{m+1} = 0) \) be a partition of size \( m + 1 \), and consider \( \varphi(\bar{r}, m+1) \).

Let \( (\psi_1, \psi_2, \ldots, \psi_h) \) be the segment decomposition of the curve \( \varphi(\bar{r}, m+1) \). We split the argument into two cases, depending on the value of \( r_m \). First suppose that \( r_m \geq \frac{k-1}{k} \). By definition, the last linear segment \( \psi_h \) of \( \varphi_{m+1} \) has the form \( \psi_h(q) = 0 \), for \( 1 - r_m \leq q \leq 1 \), and so

\[
L(\varphi_{m+1}) \geq L(\psi_h) = \sqrt{kr_m} \geq \sqrt{k-1}
\]

and we are done. We therefore assume that

\[
r_m < \frac{k-1}{k}.
\]

Consider the line \( \ell \) given by the equation \( r(q) = -kq + k(1 - r_m) \) which passes through the point \( E = (q_E, r_E) = (1 - r_m, 0) \) and has slope \(-k\). Let \( j = j(\bar{r}) \leq m \) be the largest index such that the line \( \ell \) intersects the bottom edge \( B_j \) of the square \( S_j \). Let \( D = (q_D, r_D) = ((1 - r_m) - r_j/k, r_j) \) be the intersection point of the line \( \ell \) with \( B_j \). By definition of the curve \( \varphi_{m+1} \), the next to last segment \( \psi_{h-1} \) coincides with the segment of the line \( \ell \) connecting the points \( D \) and \( E \). (see figure 4.) Let \( C = (q_C, r_C) = (1 - \frac{k-1}{k} r_m, \frac{k-1}{k} r_m) \) be the intersection point of \( \ell \) with the anti-diagonal. Note that since the squares \( S_i \) cover the anti-diagonal, the \( q \) coordinate of the point \( D \) is no
smaller than that of $C$, i.e.,

$$q_C = 1 - \frac{k}{k-1} r_m \leq q_D \leq 1 - r_m$$

Moreover, the $r$ coordinate of $D$, $r_D$, is no larger than the $r$ coordinate of $C$, $r_C$, i.e.,

$$(62) \quad r_m \leq r_j = r_D \leq r_C = \frac{k}{k-1} r_m.$$ 

Let $\tilde{\varphi}_{m+1} = \tilde{\varphi}_{m+1}(\tilde{r})$ be the part of $\varphi_{m+1}$ consisting of $\psi_1, \ldots, \psi_{h-2}$, i.e., the curve $\varphi_{m+1}$ restricted to the range $[0, q_D]$. By (61) and (62), this range is not empty. Since $\psi_{h-1}$ is a segment of the second type, $L(\psi_{h-1}) = 0$. 

**Figure 4.** The line $\ell$ contains the segment $\psi_{h-1}$, and intersects the bottom edge $B_j$ of the square $S_j$. The squares of the partition, the anti-diagonal and the line $\ell$ are all depicted by solid lines, and the dotted line is used to connect the point $D$ with its projection on the axis $q$. 

$$q_D = (1 - r_m) - r_j/k$$
Figure 5. The piecewise linear curve $GHID$ is $\tilde{\varphi}_{m+1}$, and the curve $GHIF$ is $\varphi_j$. The segment $DF$ is $\gamma$.

Consequently,

$$L(\varphi_{m+1}) = L(\tilde{\varphi}_{m+1}) + L(\psi_{h-1}) + L(\psi_h) = L(\tilde{\varphi}_{m+1}) + L(\psi_h)$$

By (60), $L(\psi_h) = \sqrt{kr_m}$.

Next, we estimate $L(\tilde{\varphi}_{m+1})$. The index $j = j(\tilde{r})$ determines the curve $\varphi_j$. Since the point $D$ lies on the bottom edge $B_j$, as noted above, the curve $\tilde{\varphi}_{m+1}$ is also the restriction of $\varphi_j$ to the range $[0, q_D]$. The curve $\varphi_j$ is defined in the domain $[0, 1 - r_j]$. Let $\gamma$ be the restriction of $\varphi_j$ to the complementary domain $[q_D, 1 - r_j] = [1 - r_m - r_j/k, 1 - r_j]$. See Figure 5 for an illustration.

By definition of the functional $L$,

$$L(\varphi_j) = L(\tilde{\varphi}_{m+1}) + L(\gamma)$$

By (59),

$$L(\gamma) = \sqrt{k} \sqrt{\frac{1}{r_{j-1} - r_j} (1 - r_j - q_D)} = \sqrt{k} \sqrt{\frac{1}{r_{j-1} - r_j} (r_m - \frac{k-1}{k}r_j)}$$

The segment $\gamma$ is contained in the bottom edge $B_j$ of the square $S_j$. The length of $\gamma$ is $r_m - \frac{k-1}{k}r_j$, and the length of $B_j$ is $r_{j-1} - r_j$. It follows that
Consequently,

\[ \partial g \bigg|_{(r_m, a)} (r_m, a) = (-a) \frac{\sqrt{k-1}}{\sqrt{2} \sqrt{1-ar_m}} - \sqrt{k} \left( \sqrt{1-\frac{k-1}{k} a} - 1 \right) \frac{1}{2 \sqrt{r_m}} \]

The equality \( \partial g \bigg|_{(r_m, a)} (r_m, a) = 0 \) holds when

\[ \sqrt{k} \left( 1 - \sqrt{1-\frac{k-1}{k} a} \right) (\sqrt{1-ar_m}) = a \sqrt{k-1} \sqrt{r_m} \]
Since $r_m < \frac{k-1}{k}$ and $1 \leq a \leq \frac{k}{k-1}$, both sides are non-negative, and thus squaring both sides results in the following equivalent equation.

$$ (64) \quad k \left( 1 - \sqrt{1 - \frac{k-1}{k} a} \right)^2 (1 - ar_m) = a^2 (k-1) r_m $$

Fix $a$ and consider (64) as an equation in the single variable $r_m$. This is clearly a linear equation. The free coefficient of this equation is positive, and thus this equation has at most one solution. Since $g_0(0, a) = g(\frac{k-1}{k}, a) = \sqrt{k-1}$ for all values of $a$, by the mean value theorem, this equation has exactly one solution. Hence the function $g_a(r_m) = g(r_m, a)$ has a unique extremum in the interval $0 \leq r_m \leq \frac{k-1}{k}$. Moreover, since $\lim_{r_m \to 0} \frac{\partial g}{\partial r_m} (0, a) = \infty$ it follows that this extremum is a maximum. Consequently, for all values of $a$, $1 \leq a \leq \frac{k}{k-1}$, and $r_m < \frac{k-1}{k}$, it holds that $g(r_m, a) \geq g(0, a) = \sqrt{k-1}$.

The equation can therefore have all values of $r_m$ as solutions, but this would imply that $g(r_m)$ is constant. While this is not the case it would prove our inequality. Another possibility is that the equation has no solutions but this would contradict the mean value theorem, therefore, the only possibility is that we have a unique solution. To complete the proof of the theorem it is enough to show that the derivative of $g_a(r_m)$ at 0 is (infinite) positive, for all $a$. Indeed the coefficient of $\frac{1}{r_m}$ in $\frac{dg_a}{dr_m}$ is positive as required. q.e.d.

11. Disk scheduling under the microscope

We have computed a first order estimate to the service time in the disk scheduling problem given as $\text{diam}(D)\sqrt{n}$, where $D$ is the space-time model associated with the request distribution $p(r, \theta)$ in theorem 25. We can ask for more accuracy, perhaps we can say something about the asymptotic statistical properties error $\text{ST}(\bar{R}) - \text{diam}(C)\sqrt{n}$. In this section we will consider this problem for the uniform density $p(r, \theta) = 1$.

To devise a strategy, let's consider a simpler example. Consider $n$ flips of an even coin and let $L_n$ be the number of heads. The first order approximation will say that $L_n$ is roughly $n/2$. The correct order of magnitude of the difference is $\sqrt{n}$. We then consider the function $F(s) = \text{Pr}(\frac{L_n - n/2}{\sqrt{n}} \leq s)$ and the central limit theorem from chapter 2 gives us a formula for $F(s)$, which turns out to be the Gaussian distribution.

Let's consider the same issues in the context of the longest chain in an interval in a $d = 2$ dimensional Minkowski space. Specifically, we consider the interval of the points $A = (0, 0)$ and $B = (1, 1)$ w.r.t. the partial order $\leq_{\text{pos}}$.

The theorem of Kerov and Vershik tells us that the longest chain $L_n$ has approximate size $2\sqrt{n}$. Taking the difference $L_n - 2\sqrt{n}$, we may ask for its order of magnitude. This is already non trivial, but it turns out that the answer is $n^{1/6}$. We may then consider the function $F(x) = \text{Pr}(\frac{L_n - 2\sqrt{n}}{n^{1/6}} \leq x)$. 

A remarkable theorem of Baik, Deift and Johansson, [21], computes the function $F(x)$. The result is the following.

**Theorem 30.** Let $q(x)$ be the unique Painlevé II function solving the differential equation

$$q''(x) = xq(x) + 2q^3(x)$$

and satisfying

$$q(x) \sim \frac{1}{2\sqrt{\pi}x^{1/4}}e^{-(2/3)x^{3/2}}$$

as $x \to \infty$. Let

$$F(x) = \exp\left(-\int_x^{\infty} (x-t)q^2(t)dt\right)$$

Where $\exp$ refers to the natural exponential function. Let $L_n$ denote the random variable describing the length of the longest increasing subsequence among $n$ uniformly chosen points in the unit square, then

$$\Pr\left(\frac{L_n - 2\sqrt{n}}{n^{1/6}} \leq x\right) \to F(x)$$

as $n \to \infty$ for all $x$.

The theorem of Baik-Deift-Johansson is a central limit theorem for longest chains, but it does not estimate rare events in which the order of magnitude of the difference $L_n - 2\sqrt{n}$ is larger than $n^{1/6}$. For that we need a large deviation result, namely an estimate for the probability of such rare events. This is provided by the following result of Lowe and Merkl, [124] which improves upon rougher estimates in [21].

**Theorem 31.** Let $0 < t_n < n^\delta$, for some $\delta < 1/3$ be a sequence such that $t_n \to \infty$. Let $l_n = t_n n^{1/6} + 2\sqrt{n}$, then

$$\lim_{n \to \infty} \frac{\ln(\Pr(L_n \geq l_n))}{-(4/3)t_n^{3/2}} = 1$$

We can use this result to prove a weak version of a theorem of K. Johansson, see [102], which tells us that an increasing sub-sequence which strays too much from the diagonal cannot be too long.

**Lemma 11.** Let $\eta > 0$. Let $\tilde{L}_{n,\eta}$ be the length of the longest chain among $P(n)$ sampled points in the unit square w.r.t. $\leq_{\text{pos}}$ which does not lie within a $n^{\eta-1/6}$ neighborhood of the diagonal, then,

$$\Pr(\tilde{L}_{n,\eta} \geq 2\sqrt{n}) < c(e^{-\delta^{3/2}})$$

for some constant $c > 0$.

**Proof:** We let $S_{i,j}$ be the square consisting of points $(x,y)$ with $i/n^{1/3} \leq x \leq (1+1)/n^{1/3}$, $j/n^{1/3} \leq y \leq (j+1)/n^{1/3}$, where $0 \leq i, j \leq n^{1/3}$, are integers.
Assume that \(|i-j| \geq n^{1/6+\eta}\) for some \(\eta > 0\). Let \(A_{i,j} = (i/n^{1/3}, j/n^{1/3})\) and \(B_{i,j} = ((i+1)/n^{1/3}, j+1/n^{1/3})\). For any point \(E \in S_{i,j}\) we have \(A_{i,j} \leq B_{i,j}\). We conclude that if a chain passes through \(S_{i,j}\) it is contained in the union of \(I^+(A_{i,j})\) and \(I_-(B_{i,j})\). Assume that \(i > j\). The area of \(I_-(B_{i,j})\) is
\[
ij/n^{2/3} = ((i+j)/2 + (i-j)/2)((i+j)/2 - (i-j)/2)n^{-2/3}
\]
\[= (1/4)((i+j)^2 - (i-j)^2)n^{-2/3} \leq (1/4)((i+j)^2 - n^{2\eta+1/3})n^{-2/3}
\]
\[\leq (1/4)(i+j)^2(1 - n^{2\eta-1/3})n^{-1/3}
\]
the last inequality because \((i+j)^2 \leq 4n^{2/3}\). The estimated number of points in a longest increasing sub-sequence in \(I_-(B_{i,j})\) is therefore at most
\[
2\sqrt{(1/4)(i+j)^2(1 - n^{2\eta-1/3})n^{1/3}} = (i+j)n^{1/6}\sqrt{1 - n^{2\eta-1/3}}
\]
\[\leq (i+j)n^{1/6} - (i+j)n^{2\eta-1/6}/4\]
the last inequality since \(\sqrt{1 - n^{2\eta-1/3}} \leq 1 - n^{2\eta-1/3}/4\) for \(n\) large enough via a simple derivative estimate. Applying theorem 31 with \(t_n = n^\eta\), the probability of having an increasing sub-sequence of size at least \(n^{1/6} - (i+j)n^{2\eta-1/6}/4 + n^{\eta}n^{1/6}\) is at most \(O(e^{-\eta^{3/2}})\). Doing the same computation for \(I^+(A_{i,j})\), we have an estimated longest increasing sub-sequence of size \((n^{1/3} - i - j + 2)n^{1/6} - (n^{1/3} - i - j + 2)n^{2\eta-1/6}/4\) and the probability of a sub-sequence of length
\[
(n^{1/3} - i - j + 2)n^{1/6} - (n^{1/3} - i - j + 2)n^{2\eta-1/6}/4 + n^{\eta}n^{1/6}
\]
is again at most \(O(e^{-\eta^{3/2}})\). Adding the two we get that the probability for a sub-sequence in the union to be at least \((n^{1/3} + 2)n^{1/6} - (n^{1/3} + 2)n^{2\eta-1/6} + 2n^{\eta+1/6} \leq 2\sqrt{n}\) is at most \(O(e^{-\eta^{3/2}})\). The sets \(s_{i,j}\) with \(|i-j| \geq n^{1/6+\eta}\) cover the set of points which is outside an \(n^{\eta-1/6}\) neighborhood of the diagonal. Adding all the probabilities over all possible \(i, j\) concludes the proof. \(q.e.d\)

Using these two results we prove the following finer estimate on \(ST(\bar{R})\).

**Theorem 32.** If \(f(\theta) = \theta\) and \(p\) is the uniform distribution then, for all \(\varepsilon > 0\), w.h.p.

\[
\left(\frac{1}{4} - \frac{1}{4}\right)(2)^{-1/6} - \varepsilon)\ln^{2/3}(n)n^{1/6} < ST(\bar{R}) - \sqrt{2n} < \left(\frac{1}{4} - \frac{1}{4}\right)(2)^{-1/6} + \varepsilon)\ln^{2/3}(n)n^{1/6}
\]

**Proof:** We will find an estimate from below for \(ST(\bar{R})\), by considering the longest sequence in a "fence" like shape. We subdivide the disk into narrow rectangular slabs \(R_t\) given by \((i-1/2)/n^{1/6-\eta} \leq t \leq (i+1/2)/n^{1/6-\eta}\). We denote the centers of the bottom edges of the slabs by \(A_t = (i/n^{1/6-\eta}, 0)\) and the centers of the top edges by \(B_t = (i/n^{1/6-\eta}, 1)\). We denote by \(H_t\), the set \(I(A_t, B_t) \cap R_t\). The \(H_t\) are very narrow hexagons, with edges given by the lines of slope 1 and −1 passing through \(A_t\) and \(B_t\) and the long edges
of $R_i$. The hexagons $H_i$ are obviously disjoint, since the $R_i$ are disjoint. Let $S_i$ denote the square diamond $I(A_i, B_i)$.

We consider $R$ as coming from a Poisson sampling of a large domain $D$ in the plane, containing the unit square. We assume that $P(\text{vol}(D)n)$ points were sampled, $R$ being the points that happened to be in the unit square.

We recall that $\leq_{\text{ver}}$ and $\leq_{\text{inc}}$ are equivalent after a rotation $Z$. Applying theorem (11) to the rotated image we see that w.h.p., the longest chain in $q$ of size $n/2$ has distribution $\text{vol}(A) \sim \mathcal{N}(\gamma, \gamma^2 n)$, where the $\text{vol}(A)$ are i.i.d. random variables with distribution $\mathcal{N}(\gamma, \gamma^2 n)$.

Let $V_{i,j}$ be the interval $I(L_i, U_j)$ of all points $R$ such that there exist $\ell \in L_i$ and $u \in U_j$ with $\ell \leq_{\text{ver}} R \leq u$. Let $M_{i,j}(m)$ be the length of the longest chain w.r.t. $\leq_{\text{ver}}$ among $P(\text{vol}(V_{i,j}))$ sampled points in $V_{i,j}$. Let $0 \leq \gamma < 1$. We will show that there exists a constant $c_\gamma > 0$ such that for all $i, j$ we have

$$Pr[L_{n/2} \geq t_n(\frac{\gamma^2 n}{2})^{1/6} + \sqrt{2n}] \geq c_\gamma Pr[M_{i,j}(\gamma^2 n) \geq t_n(\frac{\gamma^2 n}{2})^{1/6} + \gamma \sqrt{2n}]$$

To see this, we consider the mid-points $a = a_{i,j} = (\frac{i + 1}{2}/n^{1/6+\delta}, 0)$ and $b = b_{i,j} = (\frac{j + 1}{2}/n^{1/6+\delta}, 1)$ of $L_i$ and $U_j$ respectively and consider the mid-point of the segment between them $o = o_{i,j} = (\frac{i + j + 1}{2n}, 1/2)$. Let $F = F_{i,j}$ be the map which scales the $\theta, r$ plane by a factor of $\gamma$ around the point $o$. Since $V_{i,j}$ is convex and $o \in V_{i,j}$ we have $F(V_{i,j}) \subseteq V_{i,j}$. Consider the point $A = (\theta(F(a)), 0)$ and the point $B = (\theta(F(b)), 1)$. For any point $p \in F(L_i)$ we have $r(p) - r(A) = \frac{1 - \gamma}{2}$ and $|\theta(p) - \theta(A)| \leq \frac{\gamma}{2} n^{-1/6-\delta}$. The same holds with $A$ replaced by $B$ and $p$ by any point $q \in F(U_j)$. We conclude that $(\frac{1 - \gamma}{2})^2/2 - \frac{1}{2} \gamma^2 n^{-1/3-2\delta} \leq \text{vol}(I_{\text{ver}}(A, p), \text{vol}(I_{\text{ver}}(q, B)) \leq (\frac{1 - \gamma}{2})^2/2$. In addition we have $\text{vol}(I_{\text{ver}}(A, B)) \leq 1/2$. We have $\text{vol}(F(V_{i,j})) = \gamma^2 \text{vol}(V_{i,j})$ and since $F$ is a scaling of the plane which preserves the partial order $\leq_{\text{ver}}$ we see that the distribution of the longest sequence in $F(V_{i,j})$ is the same as that of the random variable $M_{i,j}(\gamma^2 n)$. Consider the event $\Omega$, that the length $\Lambda$ of the longest chain in $F(V_{i,j})$ satisfies $\Lambda \geq t_n(\frac{\gamma^2 n}{2})^{1/6} + \gamma \sqrt{2n}$. Let $\lambda_1, \lambda_\Lambda$ be the endpoints of the chain. Let $p \in L_i$ be any point such that $p \leq_{\text{ver}} \lambda_1$ and $q \in U_j$ any point such that $\lambda_\Lambda \leq_{\text{ver}} q$. Since $F(V_{i,j}) \subseteq V_{i,j}$ these points exist by the definition of $V_{i,j}$ as an interval. Consider $\psi$, the length
of the longest chain in $I_{ver}(A, p)$. Since the interior of this set is disjoint from $F(V_i,j)$ this random variable is independent of the random variable $\Lambda$. Similarly $\phi$, the length of longest chain in $I_{ver}(q, B)$, is independent of $\Lambda$.

Let $m = \left( \left( \frac{1-\gamma}{2} \right)^2 \frac{n}{2} - \frac{1}{8} \gamma^2 n^{2/3-2\delta} \right)$. The number of points sampled from $I_{ver}(A, p)$ is $P(\mu)$ with $\mu \geq m$ and similarly, the number of points sampled from $I_{ver}(q, B)$ is $P(\kappa)$ with $\kappa \geq m$. We have for $n$ large enough that $\frac{1-\gamma}{2}\sqrt{n} \leq \sqrt{2m} + m^{1/6}$ and so $Pr(L_m \geq \frac{1-\gamma}{2}\sqrt{n}) \geq F(1)$, where $F$ is the function from theorem 30. Setting $c = (F(1)/2)^2$ we get the required result.

Now letting $\gamma \to 1$, $\delta \to 0$ and recalling that if $|i - j| \geq n^{2\delta}$, there is an exponentially small probability of having a chain of size $\sqrt{2n}$ in $V_i,j$ we see by the union bound that w.h.p. we get the desired upper bound. q.e.d.

12. Lorentzian geometry and general relativity theory

As we noted earlier, Lorentzian geometry was not invented to model disk scheduling nor airplane boarding, but rather relativity theory. The universe is modeled by a 4 dimensional space-time. In his general theory of relativity, Einstein gave a prescription for relating the energy and momentum of objects in the universe to the Lorentzian metric on the modeling space-time. The equations are known as the Einstein field equations. The equations are very complicated.

Given a space-time manifold, the trajectories of point-like objects trace causal curves. Massless particles such as light travel along light-like curves, while particles with mass will travel along time-like curves. The length of a curve describes the elapsed time that will be recorded by a clock attached to a particle whose trajectory is described by the curve. It is also known as proper time. The points of the space-time are called events. An event $A$ can influence another event $B$ if and only if $A \leq_M B$. This restriction can be interpreted as saying that information or influence can only travel at or under the speed of light. Properties like causality or strong causality are considered useful since in their absence time travel to the past is possible.

We have the following notion that generalizes the notion of a straight line.

**Definition 73.** A parametrized causal curve $\gamma$ is called a geodesic if we can cover $[a, b]$ by a finite number of open sets $U_i$ such that for any pair $s_1, s_2 \in U_i$, the curve $\gamma$ restricted to the interval $[s_1, s_2]$ has maximal length among all causal curves with endpoints $\gamma(s_1)$ and $\gamma(s_2)$. In addition, we require that $\gamma$ is parametrized by arc-length.

It can be shown that geodesics are given as solutions to a system of ordinary differential equations. Using the theory of such equations it can be shown that locally, every point $x \in S$ has an open neighborhood $N_x$ such that for any $x \in N_x$ there is a unique geodesic with endpoints $x$ and $\tilde{x}$. It is also clear that a maximal curve in a domain $D$, which has no intersection with the boundary of $D$, except the endpoints is a geodesic.
A central property that relates the space-time model to gravity is Einstein’s law of motion which states that objects under the influence of gravity alone (free falling objects), will move along geodesics. We note that theorem 19 makes the interesting claim that, locally, a uniformly random particle trajectory (time-like curve) between two events closely follows a geodesic. In other words, we get a probabilistic interpretation to Einstein’s law. The free falling trajectory need not maximize proper time (an optimization problem), rather, it is completely random. I am not certain Einstein would have liked such an interpretation, after all, he did not like quantum mechanics because “god does not play dice”.

Notes for chapter 3

(1) The amazing part about theorem 30 is that the function $F(x)$ had already appeared in the asymptotic analysis of a seemingly unrelated random variable. Recall that a Hermitian matrix $H$ is a matrix with complex entries which satisfies $H = H^t$, where $H^t$ stands for the matrix where all the entries of $H$ have been complex conjugated. The equality obviously implies that the below diagonal entries of $H$ are determined by the above diagonal entries and that the diagonal entries must be real. Via its entries, we can think of the set of Hermitian matrices as an $n^2$ dimensional Euclidean space. Let $dM$ denote the standard Lebesgue measure on that space. We are interested in constructing random Hermitian matrices. We think of a complex number as consisting of two real numbers. To draw a random hermitian matrix, we draw the real and complex parts of its above diagonal entries using independent random variables with Gaussian distribution. Similarly we use more i.i.d. Gaussians to draw the real diagonal entries. The density on the space of matrices, is given by the product of the densities of all entries, which is easily seen to be $e^{-\text{tr}(H^2)}dM$, where $\text{tr}$ stands for the trace, the sum of the diagonal entries. This formula has the added feature of being invariant under conjugation $H \rightarrow UHU^{-1}$ by a unitary matrix, since the trace which is also the sum of the eigenvalues, is invariant. In fact it can be shown, [132], that the form of the density is nearly determined by the requirements of being a product of entry densities and being invariant.

Given this natural measure on the Hermitian matrices, the eigenvalues of the matrices which are real, become interesting random variables. Assume that the eigenvalues are sorted so that $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. We consider the properties of the largest eigenvalue $\lambda_n$. It turns out that $\lambda_n$ is approximately $\sqrt{2n}$ and that the order of magnitude of the error is $n^{-1/6}$. It was proven by Tracy and Widom that $Pr((\lambda_n - \sqrt{2n})2^{1/2}N^{1/6} \leq x) = F(x)$, see [171].
(2) We have considered maximal chains and uniformly random chains in discrete space-time as a tool for understanding the wave front process on such posets. However, the wave front process also involves by its definition anti-chains which form the wave front at every stage. In the disk scheduling problem these fronts correspond to requests which are serviced in the same disk rotation while in the airplane boarding problem they correspond to passengers who can sit down at the same time. Considering a domain $D$ in a strongly causal space-time, we can consider for each point $x \in D$ the function $\ell(x)$ which is given by the length of the longest causal curve in $D$ ending at $x$. Assume that $D$ is a nice domain. The longest causal curve ending at $x$ can be extended to begin at a minimal point in $D$. Letting $D_{\text{min}}$ denote the set of minimal points in $D$, we can think of the level hyper-surfaces $\ell = r$ as ”spheres” of Lorentzian radius $r$ around $D_{\text{min}}$, especially if $D_{\text{min}}$ is a single point as in our applications. The level hyper-surfaces by definition must be weak anti-chains, i.e., if both $x, y$ belong to the level hyper-surface $\ell = r$ then we cannot have $x \prec_S y$ nor $y \prec_S x$. In the same way that chains were locally Lifschitz functions from a time parameter to space parameters, anti-chains are locally Lifschitz functions from space parameters to a time parameter. The Lifschitz property tells us by a result of Radamacher, that the surface is differentiable almost everywhere. Consider a point $x$ with $\ell(x) = r + dr$. Consider a maximal length curve $\gamma_x$ ending at $x$. The curve $\gamma$ will meet the level hyper-surface $\ell = r$ at some point $y$. If the level hyper-surface happens to be differentiable at $y$ (and this happens a.e.), with tangent hyper-plane $H_y$, then necessarily, $\gamma'(y) = v_y$ will be $M_y$ orthogonal to $H_y$. The reason is the same as the reason why radial lines are orthogonal to spheres, if they were not orthogonal, we could find an even longer curve ending at $x$. Consequently, the volume between the level surfaces $\ell = r$ and $\ell = r + dr$ is $A(\ell = r)dr + o(dr)$, where $A$ is the Lorentzian area of the level hyper-surface. Consequently, the number of passengers per stage who sit down during any interval $[r, r + dr]$ of the process is asymptotically proportional to the Lorentzian area of the level hyper-surface.

(3) We considered the problem in which we are given $n$ I/Os and we wish to estimate the number of disk rotations required to service all of them. Such a scenario is sometimes referred to as batch scheduling where the $n$ requests form the batch. Another scenario which may be considered, is the situation where we are given an initial set of $n$ requests, sampled according to some density function $p(r, \theta)$ and each time that we service a request, we are given a new request, which is sampled using the same density. This way at any stage of the process we have $n$ requests outstanding. We would
be interested in the long term statistics of this process. We may consider this setting as an online setting in which decisions have to be made given a current state of information. We note however, that this setting is very different from another online setting which has been proposed and studied in [44].

The natural question in our online setting would be to find a policy that on average services the largest number of requests per disk rotation. The accompanying question is to estimate the average number of requests serviced per rotation by an optimal stochastic algorithm.

To set our expectations, we may consider the situation in which there are \( n \) points sampled w.r.t. \( p(r, \theta) \). Using the results of the present chapter we know that w.h.p the maximal number of requests that can be serviced in a single rotation, \( M \), satisfies 
\[
|M - L(p)| < \varepsilon \sqrt{n}
\]
for all \( \varepsilon > 0 \), where \( L \) is the maximal length (proper time) of a time-like curve on the unit square equipped with the Lorentzian metric 
\[
2p(r, \theta)(d\theta^2 - dr^2).
\]

Since at any given time the algorithm can choose among \( n \) points and all points were sampled w.r.t. \( p \), we expect that on average the algorithm will not be able to service more than
\[
(L(p) + \varepsilon)\sqrt{n}
\]
points per rotation. However, this is not obvious, since depending on the type of algorithm employed, the \( n \) points that the algorithm observes at any given time may be very different in their distribution from \( n \) points sampled independently using the density \( p \). It would thus be interesting to establish the above bound (or provide a counter example). Applying to the uniform distribution we have 
\[
L(p) = \sqrt{2}.
\]
It can be shown in this case that the average number of points that any algorithm can service per rotation is at most 
\[
(2\sqrt{2} + \varepsilon)\sqrt{n}.
\]

Continuing with the uniform distribution we may consider several algorithms. The simplest algorithm treats the initial \( n \) points as a batch, servicing them using the optimal ABZ algorithm, while ignoring any new requests. After servicing the original \( n \) points, the algorithm will consider the \( n \) new requests as a second batch and so on. Since as we have shown it takes roughly \( \sqrt{2n} \) rounds to service \( n \) batched requests the average service rate is roughly 
\[
\frac{1}{\sqrt{2}}\sqrt{n}
\]
requests per disk rotation.

The following algorithm does better. Divide the disk into \( k \) regions \( D_i, i = 1, ..., k \) of the form
\[
D_i = \left\{ (r, \theta) \mid \frac{i-1}{k} \leq r < \frac{i}{k} \right\}
\]
The algorithm, first considers the requests in $D_1$ as a batch and services them using the ABZ algorithm, then it considers the requests in $D_2$ and services them using the ABZ algorithm, continuing with this process for a very long time. Then, increase $k$, say by one, and repeat for an even longer period of time and so on. The average number of requests serviced by this algorithm per disk rotation is at least $(1 - \varepsilon)\sqrt{n}$ for any $\varepsilon > 0$, in fact, for any $\varepsilon > 0$ one can choose, $k$ large enough so that applying the algorithm with $k$ fixed (without increasing the value) the algorithm will achieve the required bound. The analysis, [18] is related to the analysis of a card game known as **Bulgarian solitaire**, [1, 146]

Numerically, it seems that the greedy algorithm which picks the request that it can service most quickly, given the current head position, performs even better. It would be interesting to have precise numerics regarding the greedy algorithm and perhaps a rigorous analysis of its performance.

There are many other algorithms which can be devised and explored either numerically or analytically, which may be even better. The interested reader is urged to further explore these problems.
CHAPTER 4

Mirrored configurations

In chapters 1-3 we considered systems with a single disk, such as perhaps resides in your desktop or laptop. In this chapter we will consider some issues that arise in systems with more disks.

One simple reason to have more than one disk drive is that disks fail. When they fail, the data which resides on them is lost. There are several methods to overcome this problem, when the system has more than one disk. The simplest and oldest method, which is still rather popular is to mirror the data. Mirroring means that every piece of data is written on two or more different storage devices, say disks. If one of the disks fails, the data can be read off the other disks. Mirroring data is an expensive solution since it requires twice or more storage space, therefore the number of copies is usually small, two being the most popular choice. We will assume that the number of copies is two.

While mirroring is motivated primarily by considerations of data availability, the fact that we have multiple copies of our data can be used to improve the performance of read I/O requests. If there is a single data copy and there are two requests for data items that happen to reside on the same disk drive, then they cannot be handled simultaneously. on the other hand, if each piece of data resides on two different disks, we may be able to service both requests at the same time by directing the requests to separate disks which contain the desired data.

In this chapter we will be concerned with the issue of cleverly placing the mirrored data on the disk drives, in a way that will optimize our chances of servicing requests simultaneously. We will model the configuration (layout) of mirrored data on disk drives using graphs. We state the optimization problem in graph theoretic terms. We introduce some basic results in the theory of branching processes and use them to prove a general upper bound on the utility of mirrored configurations. We then introduce some basic results concerning number theory and in particular, quaternions, to construct some optimal data placement configurations.

1. Modeling a mirrored data configuration using graphs

Assume that each disk in the system contains precisely \( d \) logical units or files, as in chapter 1. Assume that we have \( n \) disks. By our assumptions there are \( nd/2 \) files (\( d \) file per disk, but two copies for each file).
DEFINITION 74. Given a placement of files on disks in a mirrored system we construct the configuration graph \( G \) as follows. The vertices of the configuration graph correspond to the disks of the system. The edges of the graph correspond to the files in the system. Each file is represented by an edge which connects the vertices of the two disks on which its copies reside.

By our assumption that each disk contains precisely \( d \) units the configuration graph is \( d \)-regular.

As we noted in chapter 1, files or units will be placed on the disk in concentric rings of the form \( r_{i-1} \leq r < r_i \). When the system is mirrored we have the option of reading from either copy of the data, so we can choose which of the two disk drives which contains the data will service the read request. We would like to use this flexibility to minimize head movement between different files on the same disk. Write requests have to be written to both disks but there are other methods to optimize them, such as asynchronous writing, which we will not discuss here for lack of mathematical interest. Consequently, we assume that there are only read requests to the system.

It is usually the case that in any given time period not all files have substantial activity. The set of active files changes over time and so we will assume that it is a random set. Consider a set of active files, \( A \). This set induces a subgraph \( G_A \) of the configuration graph \( G \) with the same vertices and with edges corresponding only to active files. If we can assign a separate disk to service the requests of each active file, then none of the disks will experience inter-file seeks. In the language of chapter 1, the seek estimate of each disk will be zero since it will have at most one active file in any given time period. In that case we say that the configuration \( G \) can support the active set \( A \). Assigning files in \( A \) in a 1-1 fashion to disks that contain them is the same as having a 1-1 orientation of the edges of \( G_A \). It is common that a disk in the storage system will fail. It is much less common that two disks fail simultaneously. In fact a double failure can result in loss of data since the two failed disks may have both copies of the same file. We would like to eliminate seeks even after the failure of a single disk corresponding to a given vertex \( v \). This means that each file in the active set has to be assigned to a disk other than \( v \). We are led to the following definitions

DEFINITION 75. We say that a configuration graph \( G \) can support an active set of files \( A \) if the graph \( G_A \) has a 1-1 orientation of its edges. We say that \( G \) can support \( A \) after disk failure if for any vertex \( v \in V[G] \) there is a 1-1 orientation of \( G_A \) such that \( v \) is not in the image of the orientation.

The following lemma characterizes the notion of support.

LEMMA 12. A system with configuration graph \( G \) can support an active set \( A \) iff each connected component of \( G_A \) is either a tree or has a unique
cycle (is uni-cyclic). Let \( v \in G \) be any vertex. \( G \) can support \( A \) after disk failure iff \( G_A \) is acyclic.

**Proof:** Obviously a 1-1 orientation cannot exist if \( G_A \) contains a component with more edges than vertices. We conclude that each connected component of \( G_A \) with \( k \) vertices has either \( k - 1 \) edges, in which case it is a tree, or it has \( k \) edges in which case it has a unique cycle. Consider a component which is a tree. If \( v \) is a vertex of the component, choose it as the root, otherwise choose one of its vertices arbitrarily to be the root. Using the canonical orientation of the rooted tree we get the desired 1-1 mapping for the component. If the component has a unique cycle, then removing the cycle edges from the component we are left with a forest with connected components, one for each vertex of the cycle. We orient the cycle using one of the two walks of length \( l(C) \) and then choose each of the cycle vertices as a root and apply the previous procedure to its corresponding tree. Obviously, removing any vertex \( v \) of a uni-cyclic component from the range of an orientation \( f \) will not allow \( f \) to be 1-1. \( q.e.d. \)

As noted previously, the active set \( A \), which changes in time, can be considered to be random. Our model for picking a random edge subset \( A \) will be the standard random subgraph model of Erdos and Renyi, introduced in [67]. In this model each edge of \( G \) is chosen independently with some fixed probability \( \rho \) which controls the expected number of active files. The set of chosen edges is the edge set of \( G_A \). We let \( G_\rho \) denote the probability space on the subgraphs of \( G \) induced by this model. If \(|A| = k\), then the probability of \( G_A \) according to this model is \( \rho^k(1 - \rho)^{d|V|/2 - k} \). This type of random graph model is very popular and has been studied extensively, see [35, 4] for example.

The ability of a system to support random active sets depends on the structure of the configuration graph \( G \) and our goal in this chapter is to construct configuration graphs \( G \) whose ability to support random active sets \( A \) is near optimal.

We now formulate the problem more carefully. As is the case with many mathematical problems it is easier to state mathematically meaningful results if we consider what happens asymptotically as the size of our storage system grows.

Consider a family \( \bar{G} = (G_n) = (V_n, E_n) \) of \( d \)-regular graphs, with \(|V_n| \to \infty\). Consider the random variable \( \chi_{n,\rho} : G_n(\rho) \to \{0, 1\} \) whose value on \( A \) is 1 iff the subgraph of \( G_n \) induced by the edge set \( A \) is acyclic.

**Definition 76.** The cycle threshold \( CT(\bar{G}) \) of the family \( \bar{G} \) is given by

\[
CT(\bar{G}) = \text{Sup} \left\{ \rho \mid \liminf_{n \to \infty} E(\chi_{n,\rho}) = 1 \right\}
\]

The cycle threshold, \( CT(\bar{G}) \), represents the maximum \( \rho \) for which no cycle is expected to occur in \( G_A \). Roughly speaking, in terms of active file
sets we expect that a large configuration built according to the graph $G_n$ will support an active set $A$ if it has up to $(1 - \varepsilon)\rho d|V_n|/2$ active files, but not if it has $(1 + \varepsilon)\rho d|V_n|/2$ active files.

Our goal is to construct families of $d$-regular graphs, $\bar{G}$ whose cycle threshold is as large as possible (optimal).

2. An upper bound on the cycle threshold

We would like first to obtain a general upper bound on the cycle threshold, that will help us set our expectations. With that goal in mind, we first consider some basic results in the theory of branching processes. We follow [90] which is a good general reference. These processes are very useful in understanding connected components in $G(\rho)$ when $G$ is a regular graph.

2.1. Branching processes.

DEFINITION 77. Let $Z = (p_i), i = 0, 1, 2, \ldots$ be a discrete probability distribution. The branching process associated with the distribution $p_i$ starts out at stage 0 with a single element. Assume that at stage $n$ of the process each of the elements gives birth to $i$ new elements with fixed probability $p_i$ independently of the other elements and then dies. We let $Z_n$ be the random variable which counts the number of elements after $n$ stages of the process. If $Z_n = 0$ then for all $k > n, Z_k = 0$ as well. In that case we say that the process died, while if $Z_n > 0$ for all $n > 0$ we will say that the process survived. Note that $Z_1 = Z$.

Branching processes can be studied using generating functions. We claim that the generating function $f_n$ corresponding to $Z_n$ is the $n$th iterate of $f = f_Z$. For example $f_2 = f(f(x))$ and $f_3 = f(f(f(x)))$. Assume inductively that it holds for $Z_n$. Given that $Z_n = k$, the generating function for $Z_{n+1}$ will be $f_n^k$, since the birth process is performed independently for each of the $k$ elements in the $n$th generation. Taking the weighted sum over the condition $Z_n = k$ we get the desired result.

We have the following classical result on the behavior of branching processes. The result tells us that branching processes behave very differently depending on whether they have average birth rates $E(Z)$ above or below 1.

THEOREM 33. If $E(Z) \leq 1$ and $p_1 \neq 1$ then with probability 1 the process will die, while if $E(Z) > 1$, there is a positive probability that the process will survive.

Proof: By definition the probability that $Z_n = 0$ is $f_n(0) = f(f_{n-1}(0))$. The probability is a non decreasing bounded function of $n$. Since $f$ and all its derivatives are power series with non negative coefficients, all the derivatives of $f$ are non-negative and increasing. If $f(0) = 0$ then the probability of the process becoming extinct is 0. We have $E(Z) \geq 1$, with equality only if $p_1 = 1$. Thus, the case $f(0) = 0$ is consistent with the theorem. We assume
We have \( f(1) = 1 \). If \( E(Z) = f'(1) > 1 \) then for small enough \( \varepsilon > 0 \) we have \( f(1 - \varepsilon) < 1 - \varepsilon \) and looking at \( f(x) - x \) we conclude that for some \( 0 < x_0 < 1 - \varepsilon \) we have \( f(x_0) = x_0 \). Since the derivative of \( f \) is strictly increasing, \( f \) can have only two fixed points \( x_0 \) and \( 1 \). Since \( x_0 > 0 \) we have inductively \( f_n(0) < f_n(x_0) = x_0 \). On the other hand the limit \( x_1 \) of \( f_n(0) \) equals the limit of \( f_{n+1}(0) = f(f_n(0)) \) so \( x_1 \) is also a fixed point which is at most \( x_0 \), hence equals \( x_0 < 1 \). we conclude that the survival probability \( 1 - x_0 \) is strictly positive. If \( f'(1) \leq 1 \) the same arguments show that \( 1 \) is the only fixed point for \( f(x) \) in the interval \( 0 \leq x \leq 1 \). q.e.d.

For any distribution with finite mean we have

\[
E(Z_n) = E(Z)^n \tag{66}
\]

since

\[
E(Z_n) = f'(1) = f'(f_{n-1}(1)) f'_{n-1}(1) = f'(1) f'_{n-1}(1) = E(Z) E(Z_{n-1})
\]

Let us consider the random variable \( W_n = Z_n/E(Z)^n \) which has mean 1.

\textbf{Lemma 13.} If \( E(Z) > 1 \) and \( E(Z^2) \) is finite, then the random variables \( W_n \) converge with probability 1 to a random variable \( W \) with mean 1.

\textbf{Proof:} If \( Z_n = j \) then the average of \( Z_{n+k} \) will be \( j E(Z)^k = E(Z)^k Z_n \). Dividing this equation by \( E(Z)^{n+k} \) we get that the average of \( W_{n+k} \) conditioned on the value of \( W_n \), \( E(W_{n+k}|W_n) \) is \( W_n \). By 207, the second moment of \( Z_n \) can be computed from \( f''_n(1) \) and \( f'_n(1) = E(Z)^n \). The expression \( f''_n(1) \) can be computed inductively using the formula \( f''_{n+1}(1) = f''(1) f'_n(1) + f''(1)(f'_n(1))^2 \). From \( E(Z_{n}^2) \) we can compute \( E(W_{n}^2) = E(Z_{n}^2)/E(Z)^{2n} \). The end result is

\[
E(W_n^2) = 1 + \frac{V(Z)}{E(Z)^2 - E(Z)} (1 - E(Z)^{-n})
\]

Combining with the formula for conditional expectation we have

\[
E((W_{n+k} - W_n)^2) = \frac{V(Z) E(Z)^{-n}}{E(Z)^2 - E(Z)} (1 - E(Z)^{-k})
\]

This means that we have convergence in the \( L_2 \) norm to a random variable \( W \). Moreover, since the series is convergent we obtain by the Borel-Cantelli lemma (see the appendix) that convergence is pointwise with probability 1. The second moment of \( W \) is the limit of second moments

\[
E(W^2) = 1 + \frac{V(Z)}{E(Z)^2 - E(Z)}
\]

and \( E(W) = 1 \). q.e.d.

From the lemma and Chebyshev’s inequality (see the appendix) we deduce that with positive probability, say \( \eta \), we have \( W \geq 1 \). Since \( Pr((W_n - W)^2 > \varepsilon) \leq E((W_n - W)^2)/\varepsilon^2 \) we have the following corollary.
Corollary 1. There exists $\eta > 0$ such that for any $\delta > 0$ and for all $n$ large enough, $W_n > 1 - \delta$ or $Z_n \geq E(Z)^n(1 - \delta)$, with probability at least $\eta > 0$.

2.2. An upper bound for the cycle threshold. Consider an infinite $d$-regular tree $T_d$ with a root $v$. The tree can be described as follows. At level 0 we have the vertex $v$. At level 1 we add $d$ vertices and connect all of them to $v$. At level 2 or higher we add $d - 1$ new vertices for each vertex $w$ of the previous level and connect them to $w$. We let $T_{d,k}$ denote the first $k$ levels and $T_d$ is defined as the union of $T_{d,k}$.

If $w$ is a vertex we let $T_{d,k}(w)$ be the subtree consisting of vertices $u$, of level at most $k$ which are connected to $v$ through $w$. Consider the branching process where each element gives birth to $1 \leq i \leq d - 1$ new elements with probability $\frac{d - 1 - i}{i(d - 1 - \rho)}(1 - \rho)^{d - 1 - i}$ which is the binomial distribution $Z = B(d - 1, \rho)$. Consider a random graph in the model $T_{d,k}(\rho)$. Let $w \neq v$ be a vertex of level $j$, then, the number of vertices in level $j + 1$ to which $w$ is connected has the same distribution $Z = B(d - 1, \rho)$. We conclude that the number of vertices of level $k$ in $T_{d,k}(w)$ is distributed like $Z_{k-j}$. This fact will relate branching processes to properties of $d$-regular graphs.

We can now prove an upper bound on the cycle threshold. The proof, which is taken from [5] is inspired by an argument in [6].

Theorem 34. Let $G$ be a family of $d$-regular graphs, where $d \geq 2$. Then $CT(G) \leq \frac{1}{d - 1}$. Moreover, the following stronger result holds. For any integer $d \geq 2$ and any real $\varepsilon > 0$, there is a finite $n_0 = n_0(d, \varepsilon)$ so that the following holds. For any $n > n_0$ and any $d$-regular graph $G$ on $n$ vertices, if $\rho = \frac{1 + \varepsilon}{d - 1}$, then the probability that the random subgraph $G(\rho)$ of $G$, is acyclic, does not exceed $\varepsilon$.

Proof: Fix $d \geq 2$ and $\varepsilon > 0$, and let $G = (V, E)$ be a $d$-regular graph on $n$ vertices. When $d = 2$ there is nothing to prove so we assume from now on that $d \geq 3$. Throughout the proof we assume, whenever needed, that $n$ is sufficiently large as a function of $d$ and $\varepsilon$. Whenever we write a term $o(1)$, we refer to a quantity that tends to zero, as $n$ tends to infinity. We will also write $f(n) = O(g(n))$ if there exists some constant $c > 0$ such that $f(n) < cg(n)$ whenever $n$ is large enough.

Define $s = \lfloor \log_d(\log n) \rfloor$, and let $S$ be a maximal (with respect to containment) collection of pairwise edge disjoint cycles of length at most $s$ in $G$. Note, first, that if $|S| \geq \log^2 n$, then, as each cycle $C \in S$ lies completely in $G(\rho)$ with probability $\rho^{-|C|} \geq \rho^s \geq 1 \log n$, the probability that $G(\rho)$ is acyclic is at most the probability it contains no member of $S$ which is bounded by

\[(1 - \frac{1}{\log n})^{|S|} < \frac{1}{n} < \varepsilon,
\]
as needed. Thus we may and will assume that $|S| < \log^2 n$. We claim that there is a set $X \subset V$ of at least, say, $\sqrt{n}$ vertices of $G$, satisfying the following.
(i) The distance between any vertex of \(X\) and any vertex of a cycle in \(S\) is at least \(4s\).
(ii) The distance between any two vertices in \(X\) is at least \(4s\).

Indeed, the total number of vertices that lie within distance \(4s\) of some cycle in \(S\) is bounded by
\[
|S|sd^4s < O(\log^7 n).
\]
We can thus pick the vertices of \(X\) one by one, always choosing a vertex that does not lie within distance \(4s\) of any of the members of \(S\), and does not lie within distance \(4s\) of any of the previously chosen members of \(X\). As long as we have chosen less than \(\sqrt{n}\) vertices of \(X\), there are still less than \(O(\log n) + \sqrt{n}d^4s \leq O(\sqrt{n}\log^4 n)\) vertices that cannot be picked, providing the required set \(X\) of size at least \(\sqrt{n}\) (with room to spare).

By the choice of \(X\) and the maximality of \(S\), each vertex \(v \in X\) does not lie within distance \(s\) of any cycle of length at most \(s\) in \(G\) (since each such cycle intersects at least one member of \(S\), and this member is far from \(v\).) Therefore, the graph restricted to vertices of distance at most \(s\) from a vertex \(v \in X\) is a \(d\)-regular tree, which we call \(T_{s,v}\).

It is convenient to consider the random subgraph \(G(\rho)\) as a union of two independently chosen random subgraphs \(G(\rho_1)\) and \(G(\rho_2)\), where \(\rho_1 = \frac{1+\varepsilon/2}{d-1}\) and \(\rho_2 \geq \frac{\varepsilon}{2d}\) is chosen such that \((1-\rho_1)(1-\rho_2) = 1-\rho\). Pick a vertex \(v \in X\) and expose all the edges of \(G(\rho_1)\) that lie in \(T_{s,v}\). By corollary 1 there is an \(\eta > 0\) such that with probability at least \(\eta\) the connected component of \(v\) in this exposed subgraph of \(G(\rho_1)\) has at least, say, \((1 + \varepsilon/10)^s\) leaves. We next show that conditioning on this event, the random graph \(G(\rho)\) will contain a cycle with probability \(1 - o(1)\). Indeed, conditioning on this event, we keep following this branching process. To do so, choose in each step a yet unexplored leaf \(u\) of the connected component, and in case there are \(d-1\) edges of \(G\) emanating from it to vertices outside this component, expose the corresponding \(d-1\) edges of \(G(\rho_1)\). If this is not the case, that is, if \(u\) is adjacent in \(G\) to at least one other vertex of the component besides its parent in the component, we declare \(u\) to be an explored vertex, and keep it as a leaf. Note that whenever we scan the edges of an unexplored vertex, the behavior is precisely as in the usual branching process on the infinite \(d\)-regular tree \(T_d\), and hence, since \(\rho > \frac{1}{d-1}\) and the number of leaves in the component grows with \(n\), the component will keep growing with high probability, and will stop with a component \(K\) containing among its leaves at least \((1 + \delta)^s\) leaves, for some \(\delta > 0\), each having at least one neighbor in \(K\) besides its parent. We can now expose the random edges of \(G(\rho_2)\), and if these edges contain any one of the edges connecting such a leaf to its neighbor in \(K\) besides its parent, we get a cycle in \(G(\rho)\). As the number of these edges grows with \(n\), this happens with probability \(1 - o(1)\), and we
conclude that indeed if the component of $v$ in $G(\rho_1)$ grows at the beginning, as assumed, then a cycle emerges with high probability.

In case the component of $v$ in $G(\rho_1)$ dies early, we pick another vertex $v' \in X$ and repeat the same process from there. Since each such vertex gives a growing component with probability at least $\eta$, and we have $\sqrt{n}$ vertices, we conclude that with probability $1 - o(1)$ there will be a vertex $w \in X$ whose component will grow, providing, w.h.p., the required cycle. q.e.d.

3. Lubotzky-Phillips-Sarnak (LPS) graphs

We have seen that the best we can hope for is a family of graphs with $CT(\overline{G}) = \frac{1}{d-1}$. In this section we prove the existence of explicit families of $d$-regular graphs with $CT(G) = \frac{1}{d-1}$, using well known number theoretic graph constructions known as LPS graphs. Moreover, currently there are no known fully explicit families of graphs with this property which are not based on the kind of number theoretic arguments that we will describe. The reader may also consult the original paper, [126], the elementary book [56], or the more difficult [125, 155], for a more thorough treatment of the construction and properties of these graphs. We begin with some number theoretic preliminaries. The subsection can be skipped by readers familiar with number theory.

3.1. Number theoretic preliminaries.

**Definition 78.** Let $p, q$ be prime numbers. We define the **quadratic residue symbol** $(\frac{p}{q})$ to be 1 if the equation $x^2 = p \pmod{q}$ has an integer solution, and $-1$ otherwise.

By definition, the symbol only depends on $p \pmod{q}$.

We state a classical result of Gauss regarding the quadratic residue symbol, see [89], [160] or [56], for a proof.

**Theorem 35.** (Quadratic reciprocity) If $p, q$ are distinct, odd, prime numbers then

$$(\frac{p}{q})(\frac{q}{p}) = (-1)^{(\frac{p-1)(q-1)}{4}}$$

Let $m, n$ be integers. We use the notation $m|n$ to denote that $m$ divides $n$. We also define Euler’s totient function

$$\phi(m) = m \prod_{p|m}(1 - \frac{1}{p})$$

which counts the number of residues prime to $m$, modulo $m$.

We will need the following classical theorem of Dirichlet, [65], about primes in arithmetic progressions.

**Theorem 36.** If $a, m$ are relatively prime numbers, i.e., they have no common divisor, then there are infinitely many primes $q$ such that $q = a$ (
mod \ m). Moreover, consider the ratio

$$r(a, m, x) = \frac{\text{Primes } q \leq x, \text{ which satisfy, } q = a \pmod{m}}{\text{Primes } q \leq x}$$

then,

$$\lim_{x \to \infty} r(a, m, x) = \frac{1}{\phi(m)}$$

Dirichlet’s remarkable proof of this theorem is provided in [59], including his famous class number formula, see also [160].

3.2. Gaussian integers and Hamiltonian integers. We consider the basic arithmetic of two systems of numbers which extend the usual integers.

Definition 79. The set of integer Hamiltonian quaternions, or Hamiltonian integers is the set of expressions of the form $$a_0 + a_1i + a_2j + a_3k$$, with $$a_i \in \mathbb{Z}$$, where addition and multiplication are given using the usual rules and the relations

$$i^2 = j^2 = k^2 = -1$$

and

$$ij = -ji = k, \ jk = -kj = i, \ ki = -ik = j$$

Hamiltonian rational quaternions are defined in the same way only assuming that $$a_i \in \mathbb{Q}$$. The Hamiltonian integers contain the Gaussian integers, which consist of elements of the form $$a_0 + a_1i$$. These can also be viewed as complex numbers with integer real and complex parts. For an element $$\alpha = a_0 + a_1i + a_2j + a_3k$$, define the conjugate

$$\bar{\alpha} = a_0 - a_1i - a_2j - a_3k$$

and the norm

$$N(\alpha) = \alpha\bar{\alpha} = a_0^2 + a_1^2 + a_2^2 + a_3^2.$$  

If $$\alpha = a_0 + a_1i + a_2j + a_3k$$ and $$\beta = b_0 + b_1i + b_2j + b_3k$$ are quaternions, then it is easy to verify that

$$N(\alpha\beta) = N(\alpha)N(\beta)$$

As quaternions with rational coefficients $$\alpha^{-1} = \bar{\alpha}/N(\alpha)$$. Consequently, a Hamiltonian integer $$\alpha$$ is invertible if and only if $$N(\alpha) = 1$$.

Definition 80. An invertible quaternion is called a unit.

An easy inspection shows that there are 8 units given by

$$1, -1, i, -i, j, -j, k, -k$$

We state some basic results about the arithmetic of the Hamiltonian integers. We also state the analogous results for Gaussian integers as we will need them later. Our goal is to establish analogues of unique factorization in these number systems. The first lemma provides a partial analogue in
Hamilton integers for division with remainder in the integers. For Gaussian integers we obtain a fully analogous result.

**Theorem 37.** If $\alpha$ is a Hamiltonian integer and $\beta$ is a Hamiltonian integer with odd norm, then there exist Hamiltonian integers $\gamma$ and $\delta$ such that $\delta = \alpha - \gamma \beta$ and $N(\delta) < N(\beta)$.

The same is true for arbitrary Gaussian integers $\alpha$ and $\beta$, with $\gamma$ and $\delta$ also being Gaussian integers.

**Proof:** We consider first the case where $\beta = m \in \mathbb{Z}$ is an odd integer. Let $a_i$ be the coordinates of $\alpha$ and let $\gamma_i$ be such that $|\alpha_i - \gamma_i m| < m/2$. This is possible since $m/2$ is not an integer. Taking $\gamma$ to have coordinates $\gamma_i$ then The norm inequality $N(\delta) < N(\beta) = m^2$ is obviously satisfied.

We now consider a general $\beta$ with odd norm. Consider $\alpha' = \alpha \overline{\beta}$ and $m = N(\beta) = \beta \overline{\beta}$. We apply the first case with $\alpha'$ and $m$. We obtain $\gamma$ and $\delta' = \alpha \overline{\beta} - \gamma m = (\alpha - \gamma \beta) \overline{\beta}$ such that

$$N(\delta') = N(\alpha - \gamma \beta) N(\overline{\beta}) < m N(\overline{\beta}) = N(\beta)^2$$

from which we deduce the result with $\delta = \alpha - \gamma \beta$ by dividing by $N(\beta) = N(\overline{\beta})$.

If $\alpha$ and $\beta$ are Gaussian, consider $\alpha/\beta = a + bi$, where $a$ and $b$ are real and let $u, v$ be integers such that $|a - u| \leq 1/2$ and $|b - v| \leq 1/2$. Let $\gamma = u + vi$ then

$$N(\delta) = N(\alpha - \gamma \beta) = N(\alpha/\beta - \gamma) N(\beta) \leq (1/2) N(\beta)$$

as required. q.e.d.

The next step is to obtain a partial analogue for the notion of greatest common divisor in Hamilton integers.

**Definition 81.** We say that $\delta$ is a right divisor of $\alpha$ if $\alpha = \gamma \delta$. We say that $\delta$ is a right greatest common divisor (r.g.c.d.) of Hamilton integers $\alpha$ and $\beta$, if $\delta$ is a right divisor of both and if any other right divisor $\gamma$ of $\alpha$ and $\beta$ is also a right divisor of $\delta$.

It is easy to check from the definition that an r.g.c.d., if it exists, must be unique up to multiplication on the right by a unit.

**Theorem 38.** Let $\alpha, \beta$ be Hamiltonian integers, $\beta$ having odd norm, then, $\alpha$ and $\beta$ have an r.g.c.d. $\delta$. Furthermore, $\delta$ can be written as $\delta = \mu \alpha + \nu \beta$ for some Hamiltonian integers $\mu, \nu$.

The same holds for arbitrary Gaussian integers $\alpha, \beta$.

**Sketch of proof:** In the following two results we just indicate how the proof goes. full details can be found in [63]. We recall that in the integers we apply the Euclidean algorithm to find the g.c.d. of two integers $\alpha, \beta \in \mathbb{Z}$. We set $\alpha = r_1, \beta = r_2$ and inductively we define $r_{k+2}$ to satisfy $r_k =
$q_k r_{k+1} + r_{k+2}$ with $|r_{k+2}| < |r_{k+1}|$. Since the sequence $|r_k|$ is a decreasing sequence of integers we will have at some point $r_{m+2} = 0$ or $r_m = q_m r_{m+1}$. Working the algorithm backwards it is not hard to show that $\delta = r_{m+1}$ is the g.c.d. and that the desired representation $\delta = \mu \alpha + \nu \beta$ exists. This proof works equally well for the Gaussian integers with the norm replacing the absolute value.

We wish to apply the same strategy for Hamiltonian integers to deduce theorem 38 from theorem 37. The main obstacle is that theorem 37 applies only when we divide by a quaternion of odd norm and the remainder may not be odd. To overcome this problem we need an additional lemma which decomposes a quaternion of even norm.

**Lemma 14.** If $\alpha$ is a Hamiltonian integer, then, $\alpha$ can be written as $\alpha = 2^k \gamma \beta$, where $N(\beta)$ is odd and $\gamma$ is one of the quaternions $1, 1 + i, i + j, 1 + k, (1 + i)(1 + j), (1 + i)(1 + k)$.

Moreover, the decomposition is unique up to units.

**Sketch of proof:** If $N(\alpha)$ is even then either all coordinates $\alpha_i$ are even or just two of them are even. In the first case $\alpha$ is divisible by 2. In the second case assume without loss of generality that $\alpha_0$ and $\alpha_1$ are even, then, it is easy to verify that $\alpha(1 - i)$ has all coordinates even and therefore $\alpha(1 - i)/2 = \alpha/(1 + i)$ is a Hamiltonian integer. In either case we can continue by induction to decompose $\alpha$, dividing it by either $2, 1 + i, 1 + j, 1 + k$ until we get a quaternion with odd norm. The lemma then follows easily by studying the products of elements of the form $1 + i, 1 + j, 1 + k$. q.e.d.

**Sketch of proof of 38 continued:** Given the lemma we can proceed according to the Euclidean algorithm replacing $r_k$ of even norm by it’s odd factor as in the lemma. It can be verified that the proof of theorem 38 from theorem 37 can be carried out along the same lines as the proof in the case of the integers. for full details see [63]. q.e.d.

**Definition 82.** Let $\alpha$ and $\beta$ be Hamiltonian integers of odd norm. We say that $\alpha, \beta$ are right sided relatively prime, if their r.g.c.d. is 1.

**Theorem 39.** Let $\alpha$ be a Hamiltonian integers and $m$ an odd integer. $\alpha$ and $m$ are right sided relatively prime if and only if $N(\alpha)$ and $m$ are relatively prime as integers.

**Proof:** If two quaternions $\alpha, \beta$ have a non trivial r.g.c.d. $\delta$, then $N(\delta)$ will be a common factor of $N(\alpha)$ and $N(\beta)$, thus we only need to prove the converse, for $\beta = m$ an odd integer. Assume that the r.g.c.d. is 1 and by theorem 38 write $\mu \alpha + \nu m = 1$. It follows that $N(\mu)N(\alpha) = N(1 - \nu m) = N(1 - \nu m) = 1 - (\nu + \bar{\nu})m + N(\nu)m^2$. We conclude that $N(\alpha)$ and $m$ are right sided relatively prime as desired. q.e.d.
Definition 83. We say that a Hamiltonian integer \( \alpha \) is irreducible if for any decomposition \( \alpha = \beta \gamma \), either \( \beta \) or \( \gamma \) is a unit.

The next result says that an irreducible (prime) integer \( p \) is never an irreducible Hamiltonian integer. To see this, we will prove that \( p \) is a sum of four squares, i.e., there exist \( x_0, x_1, x_2, x_3 \in \mathbb{Z} \) such that \( x_0^2 + x_1^2 + x_2^2 + x_3^2 = p \). We deduce that \( \alpha = x_0 + x_1i + x_2j + x_3k \) satisfies \( \alpha \bar{\alpha} = p \), hence \( p \) is not irreducible. We will later prove a stronger result, however, its nice to give a simple proof of this famous result of Legendre.

Theorem 40. (Legendre) Every prime integer \( p \) is the sum of four squares.

Proof: Following Euler, we claim that there are integers \( x, y \) such that \( 1 + x^2 + y^2 = 0 \mod p \). If \(-1\) is a quadratic residue modulo \( p \), then we can find \( x \) such that \( x^2 = -1 \mod p \) and we can take \( y = 0 \). Otherwise, consider the sequence \( p - 1, p - 2, ..., 1 \) and let \( a \neq p - 1 \) be the first element in the sequence which is a quadratic residue modulo \( p \). Let \( b = a + 1 \), then \( b \) is a non quadratic residue. The product of non quadratic residues is a quadratic residue. This follows from the fact that there are equal amounts of quadratic residues and non quadratic residues and the fact that the product of quadratic residues is quadratic. We see that \(-b\) is a quadratic residue. Letting \( x \) be such that \( x^2 = a \mod p \) and \( y \) be such that \( y^2 = -b \mod p \) finishes Euler’s argument.

By theorem 39, the quaternions \( \alpha = 1 + xi + yj \) and \( \bar{\alpha} \) have a non trivial r.g.c.d., \( \delta \). We cannot have \( \delta = p \) since \( p \) does not divide \( \alpha \). We conclude that \( \delta \) is a non trivial right divisor of \( p \). Since \( N(\delta) \neq 1 \), and \( N(\delta) \neq N(p) = p^2 \) we must have \( N(\delta) = p \), or \( \delta \bar{\delta} = p \) as required. q.e.d.

The analogous result for Gaussian integers is the following.

Theorem 41. Let \( p \) be a prime integer. If \( p = 1 \mod 4 \), then \( p \) is of the form \( \alpha \bar{\alpha} \) and \( \alpha \) is unique up to multiplication by a unit. If \( p = 3 \mod 4 \) then \( p \) is an irreducible Gaussian integer.

Proof: If \( p = 3 \mod 4 \) then it cannot be a norm of a Gaussian integer since the square of an even number is divisible by 4 and the square of an odd number is 1 \( \mod 4 \). If \( p = 1 \mod 4 \) then as we have seen \(-1\) is a square residue of \( p \) and hence there is an \( x \) such that \( 1 + x^2 = 0 \mod p \) and we can proceed as in the previous theorem. To see uniqueness, assume that \( p = \alpha \bar{\alpha} = \beta \bar{\beta} \), then \( \alpha \) and \( \beta \) are irreducible. If they are relatively prime to each other, i.e., their g.c.d. is 1, then we can write \( 1 = \mu \alpha + \nu \beta \) or \( \bar{\beta} = \mu \alpha \bar{\beta} + \nu \beta \bar{\beta} \). By commutativity we see that \( \alpha \) divides both sides so up to a unit it equals \( \bar{\beta} \), the other option is that \( \alpha \) and \( \beta \) are not relatively prime, but then since their norm is prime they again differ by a unit. q.e.d.

The last few results lead to the following characterization of irreducible Hamiltonian integers.
Theorem 42. A Hamiltonian integer $\alpha$ is irreducible iff $N(\alpha)$ is prime.

Up to multiplication by units, a Gaussian integer $\alpha$ is irreducible if $N(\alpha) = 1 \pmod{4}$ is a prime, or $\alpha$ is an integer prime $p \in \mathbb{Z}$ with $p = 3 \pmod{4}$, or $\alpha = 1 + i$.

Proof: If $N(\alpha)$ is prime then obviously it has no non trivial divisors. For the converse, assume that $N(\alpha)$ is not a prime and let $p$ be some integer prime divisor of $N(\alpha)$. By theorem 39, $\alpha$ and $p$ have a non trivial r.g.c.d.

Assuming $\alpha$ is an irreducible quaternion, we conclude that it divides $p$. It has to be a non trivial factor since otherwise $p$ itself would have been an irreducible quaternion in contradiction with the previous theorems. We conclude that $N(\alpha) = p$ a contradiction. The same argument yields the result for Gaussian integers. q.e.d.

3.3. Counting Gaussian and Hamiltonian integers of a given norm. We will need formulas for the number of Hamiltonian and Gaussian integers of a given norm $n$. Luckily, there are exact formulas which are due to Jacobi. We start with the Gaussian integers.

Consider the function $\chi(n) = \sin(\frac{\pi n}{2})$ which has period 4. We can also define $\chi$ via the formula $\chi(n) = (-1)^{(n-1)/2}$ when $n$ is odd and $\chi(n) = 0$ for $n$ even. The function $\chi$ is multiplicative, namely, for any $n,m \chi(nm) = \chi(n)\chi(m)$.

In this and the next result we follow closely [89], theorems 278, 315 and 338.

Theorem 43. Let $n$ be an integer and write $n$ as

$$n = 2^a \prod p_i^{r_i} \prod q_j^{s_j} = 2^a n_1 n_3$$

where $p_i$ are distinct primes equal to 1 modulo 4 and $q_j$ are distinct primes which are equal to 3 modulo 4. The number of Gaussian integers of norm $n$ is 0 unless $n_3$ is a square. If $n_3$ is a square the number of Gaussian integers of norm $n$ is $4d(n_1)$, where $d(k)$ is the number of divisors of $k$. Alternatively, the number of Gaussian integers of norm $n$ is $\sum_{d|n} \chi(d)$.

Proof: We first note that $2 = (1+i)(1-i) = -i(1+i)^2$. Let $\alpha = a_0 + a_1 i$ be a Gaussian integer of norm $n$. Decomposing $\alpha$ into irreducible elements and using theorem 42 we have

$$\alpha = u(1+i)^{2a} \prod \pi_i^{r_{i,1}} \prod q_j^{s_j}$$

where $\pi_i$ are such that $N(\pi_i) = p_i$ and for all $i, r_{i,1} + r_{i,2} = r_i$. We see that all $s_j$ must be even, or equivalently that $n_3$ is a square. There are 4 choices for the unit $u = 1, -1, i, -i$ and $\Pi_i(r_{i,1} + 1)$ choices for the $r_{i,1}$, but that is precisely the number of choices for a divisor of $n_3$, we simply need to choose the multiplicity of any prime divisor of $n_3$. 


To prove the second characterization note that
\[ \sum_{d|n} \chi(d) = \sum_{d|n_1n_3} \chi(d) \]
since \( \chi(m) = 0 \) for any even \( m \). The divisors of \( n_1n_3 \) correspond to the summands of the product
\[ \Pi_i(1 + p_i + ... + p_i^{r_i})\Pi_j(1 + q_j + ... + q_j^{s_j}) \]
By the multiplicativity of \( \chi \) we have
\[ \sum_{d|n_1n_3} \chi(d) = \Pi_i(\chi(1) + \chi(p_i) + ... + \chi(p_i)^{r_i})\Pi_j(\chi(1) + \chi(q_j) + ... + \chi(q_j)^{s_j}) \]
By definition and the 4 periodicity of \( \chi \) we have \( \chi(p_i) = 1 \) while \( \chi(q_i) = -1 \). If \( s_j \) is odd then the corresponding sum is 0, while if \( s_j \) is even then the sum is 1. For the \( p_i \) we get \( r_i + 1 \) so the product evaluates to \( d(n_1) \) if \( n_3 \) is a square and zero otherwise as stated before. \textit{q.e.d.}

The formula for the number of Hamiltonian integers of a given norm is given in the following theorem which is proved in the appendix.

Theorem 44. \textit{(Jacobi’s formula)} The number of Hamiltonian integers of norm \( n \) is
\[ 8 \sum_{d'|n} d' \]
where \( d' \) denotes divisors of \( n \) which are not divisible by 4.

In particular, if \( n \) is odd then the number of Hamiltonian integers of norm \( n \) is given by
\[ 8 \sum_{d|n} d \]

3.4. The LPS construction. We describe a well known family of graphs, the bipartite LPS Ramanujan graphs, constructed by Lubotzky, Philips and Sarnak in [126]. We will later show that these families have optimal cycle thresholds. We will provide two distinct points of view for the construction of these graphs. the first as a Cayley graph of a group of matrices. The second construction will be in terms of Hamiltonian integers. We follow closely the exposition in the first three sections of [126].

3.4.1. Construction of LPS graphs via Hamiltonian integers. Given \( p \) let \( \Lambda_p'(2) \) be the set of quaternions with \( a_0 \) odd, \( a_1,a_2,a_3 \) even and with \( N(\alpha) = p^k \) for some integer \( k \). It is easy to verify that \( \Lambda_p'(2) \) is closed under multiplication and conjugation.

Lemma 15. There are \( p + 1 \) elements in \( \Lambda_p'(2) \) which have norm \( p \).
**Proof:** By theorem 44 there are \(8(p + 1)\) Hamiltonian integers of norm \(p\). Given such an element \(\alpha = x_0 + x_1i + x_2j + x_3k\), we claim that there is a unique unit \(u\) such that \(u\alpha \in \Lambda'_p(2)\).

To see this claim, we notice that \(p\) being odd means that either one or three of the \(x_i\) are odd. We also have that if \(x_i\) is odd then, \(x_i^2 \pmod{4} = 1\), hence if three of them were odd then that would imply that \(p \pmod{4} = 3\). We conclude that only one of the \(x_i\) is odd. By multiplying by the appropriate (unique) unit we can move the odd \(x_i\) to the constant position and flip sign if needed. Since there are 8 units, we obtain the result. \(q.e.d.\)

We note that if \(\alpha\) has norm \(p\) then obviously \(\bar{\alpha}\) also has the same norm.

Let \(\alpha_m = a_{0,m} + a_{1,m}i + a_{2,m}j + a_{3,m}k, m = 1, \ldots, p + 1\), be the set of Hamiltonian integers in \(\Lambda'_p(2)\) of norm \(p\). We assume that the \(\alpha_m\) are indexed so that \(\alpha_{m+(p+1)/2} = \bar{\alpha}_m\) for \(m = 1, \ldots, (p+1)/2\).

**Definition 84.** Let \(S\) be the set of all \(\alpha_i\). A word of length \(k\) in the elements of \(S\) is the product of elements in a sequence \(\alpha_{i,1}, \ldots, \alpha_{i,k}\). We will say that the word is reduced if there are no pairs of consecutive elements which are conjugate, i.e., there are no pairs \(\alpha_j\bar{\alpha}_j\) appearing in the word.

We have the following unique factorization type theorem for Hamiltonian integers with norm \(p^k\), in terms of the elements in \(S\).

**Theorem 45.** (Gerritzen, van der Put [79]) Let \(\alpha\) be a Hamiltonian integer with \(N(\alpha) = p^k\), then \(\alpha\) can be uniquely written as

\[
\alpha = up^r w
\]

where \(u\) is a unit and \(w\) is a reduced word in \(S\) of length \(k - 2r\).

Consequently, if \(\alpha \in \Lambda'_p(2)\) then it can be written uniquely as \(\alpha = vp^r w\), where \(v = 1\) or \(v = -1\), and \(w\) is a reduced word in the \(\alpha_i\).

**Proof:** We use induction to show that the decomposition exists. If \(k = 1\) then there is nothing to prove since we have shown that all integer quaternions of norm \(p\) are of the form \(u\alpha_i\) where \(u\) is a unit. More generally, if \(k > 1\), we know by theorem 42 that it is not a prime quaternion and thus can be written as \(\alpha = \beta\gamma\), where \(\beta\) and \(\gamma\) have smaller \(p\) power norms. Applying this inductively we may assume that \(N(\gamma) = p\) and so \(\alpha = \beta u'\alpha_i\) for some unit \(u'\) and \(\alpha_i \in S\). Applying induction on \(\beta u'\) we get that \(\alpha = u\alpha_{i,1}\alpha_{i,2} \cdots \alpha_{i,k}\). The sequence \(\alpha_{i,j}\) may not be reduced. The product of any consecutive pair which satisfies \(\alpha_{i,j} = \bar{\alpha}_{i,j+1}\) is \(p\), which commutes with all Hamiltonian integers and hence can be moved up front. Repeating this process until the remaining word is reduced yields the desired representation.

To show uniqueness we count the number of elements on each side. By Jacobi’s formula the number of elements of norm \(p^k\) is

\[
8 \sum_{d|p^k} d = 8 \sum_{i=0}^{k} p^i = 8 \frac{p^{k+1} - 1}{p - 1}
\]
The number of reduced words of a given length $m$ is 1 if $m = 0$ and 
$(p + 1)p^{m-1} = p^m + p^{m-1}$ otherwise, since we have $p + 1$ choices for the 
first entry and each sequence element disqualifies its conjugate from being 
the next element. Summing over all possible $0 \leq r \leq k/2$ yields equality 
and therefore unique representation. The second statement of the theorem 
follows immediately since the $\alpha_i$ belong to $\Lambda'_p(2)$ and the only units in that 
multiplicative set are 1, $-1$. q.e.d.

**Definition 85.** Let $\Lambda_p(2)$ be the group obtained by identifying elements 
$\alpha, \beta \in \Lambda'_p(2)$ whenever $\alpha = (-1)^i p^j \beta$ for some $i, j$. We denote by $[\alpha]$ the 
resulting class of $\alpha$.

Let $\alpha \in \Lambda'_p(2)$. We say that $\alpha$ is a normalized element if $\alpha$ satisfies, 
a_0 > 0 is odd, $a_1, a_2, a_3$ are even and not all are $a_i$ divisible by $p$.

Given the equivalence relation, any element in $\Lambda_p(2)$ can be represented 
as a class $[\alpha]$ for a unique normalized element $\alpha$. Note also that $[\alpha]^{-1} = [\bar{\alpha}]$.

Consider the Cayley graph of the group $\Lambda_p(2)$ with generators 
$[\alpha_1], ..., [\alpha_{p+1}]$

We claim that theorem 45 states that it is an infinite $p + 1$-regular tree. To see this, note that reduced words correspond to non backtracking walks in 
the graph. A cycle in the graph would have produced a non-trivial representation 
of the element 1, in contradiction with uniqueness. In the language 
of group theory we say that the $\alpha_i$ freely generate $\Lambda_p(2)$.

**Definition 86.** Let $q \neq p$ be a prime. The subgroup $\Gamma_q \subset \Lambda_p(2)$ is 
defined by classes of elements $\alpha \in \Lambda'_p(2)$, for which $a_1, a_2, a_3$ are all divisible 
by $2q$.

The group $\Gamma_q$ is a normal subgroup. This can be seen by considering 
$H_{2q}$, the ring of quaternions with coefficients in $\mathbb{Z}/2q\mathbb{Z}$. The elements of $H_{2q}$ 
are of the form $a_0 + a_1i + a_2j + a_3k$, where $a_i \in \mathbb{Z}/2q\mathbb{Z}$ and $i, j, k$ satisfy the 
same multiplicative relations as in the Hamiltonian integers. We let $H_{2q}^*$ be 
the invertible elements in the ring. The center of $H_{2q}^*$ consists of the group 
$\mathbb{Z}$ of quaternions with $a_0 \neq 0$ and $a_1, a_2, a_3 = 0$. Taking the residues of the 
coefficients modulo $2q$ defines a map of $\Lambda'_p(2)$ into $H_{2q}$. The image lies in 
$H_{2q}^*$ since the norm of an element in $\Lambda'_p(2)$ is invertible modulo $2q$. The map 
also induces a group homomorphism $\Lambda_p(2) \mapsto H_{2q}^*/\mathbb{Z}$. The group $\Gamma_q$ is the 
kernel of this map, which shows that it is normal.

**Definition 87.** The graph $X_{p,q}$ is the Cayley graph of the quotient group 
$\Lambda_p(2)/\Gamma_q$ with respect to the generating set consisting of the images of the 
elements $\alpha_i \in S$.

The group mapping $\Lambda_p(2) \mapsto \Lambda_p(2)/\Gamma_p$ induces a graph homomorphism 
from the infinite $p + 1$ regular tree to $X_{p,q}$. Walks of length $k$ in the graph 
$X_{p,q}$, beginning and ending at the identity element, are the images of walks
in the $p + 1$ regular tree, starting at the identity and ending at an element of $\Gamma_q$. By the discussion above such walks correspond in turn to products $[\alpha_{i_1}] \cdot \ldots \cdot [\alpha_{i_k}] \in \Gamma_q$. The assumption that no backtracking is allowed in a walk and the freeness of the generators $[\alpha_i]$ means that $\alpha = \alpha_{i_1} \cdot \ldots \cdot \alpha_{i_k}$ is normalized up to its sign and we have $N(\alpha) = p^k$.

Let $W_v$ denote the number of such walks in $X_{p,q}$ which start and end at a vertex $v$. Since $X_{p,q}$ is vertex transitive we conclude that for each $v$, $W_v$ is equal to the number of integer solutions of the equation

$$x_0^2 + q^2(x_1^2 + x_2^2 + x_3^2) = p^k$$

where $x_0 + x_1i + x_2j + x_3k$ is a normalized element. This fact, which is implicit in [126], is the key to our computations since it allows us to estimate the number of cycles.

**Definition 88.** A solution $x_0, x_1, x_2, x_3$ to equation (67) such that $x_0 + x_1i + x_2j + x_3k$ is a normalized element is called a normalized solution.

3.4.2. A description of $X_{p,q}$ via matrices. We claim that if $q = 1$ (mod 4) there exists an element $i \in F_q$ which satisfies $i^2 = -1$ (mod $q$). If $g \in F_q^*$ is a generator of the cyclic group, we can take $i = g^{(q-1)/4}$, which is defined since $q = 1$ (mod 4). The element will satisfy the requirements since $i^4 = 1$ but $i^2 \neq 1$.

**Definition 89.** Let $PGL_2(q)$, the projective linear group, denote the multiplicative group of invertible 2 by 2 matrices with elements in $F_q$, where we identify two matrices $M_1, M_2$ if $M_1 = cM_2$ for some $c \in F_q^*$.

This group has $q^2 - q$ elements. This can be seen by noting that there are $(q^2 - 1)(q^2 - q)$ invertible matrices (the first row has to be non-zero, while the second, not on the line defined by the first). The equivalence class of an invertible matrix contains precisely it’s $q - 1$ non zero multiples.

Let $a = a_0 + a_1i + a_2j + a_3k$ be an element in $\Lambda_0'(2)$ of norm $p$. Let $\gamma(a)$ be the matrix

$$\begin{pmatrix} a_0 + a_1i & a_2 + a_3i \\ -a_2 + a_3i & a_0 - a_1i \end{pmatrix}$$

where the coefficients $a_m$ are reduced modulo $q$ and $i \in F_q$, the element defined above. It is easy to check that $det(\gamma(a)) = N(a)$, hence $\gamma(a)$ will be invertible. We consider $\gamma(a)$ as an element in $PGL_2(q)$, this allows us to consider $\gamma(a)$ as a mapping from $\Lambda_p'(2)$. The subgroup $\Gamma_q$ is the kernel of the mapping and we obtain an alternative description of $X_{p,q}$ as the connected component of the identity in the Cayley graph of $PGL_2(q)$ w.r.t. the images $\gamma(\alpha_k)$. It can be shown, [126], that this component is all of $PGL_2(q)$ but we will not need this fact.

We conclude from this description that

$$|V(X_{p,q})| \leq |PGL_2(q)| = q^3 - q \leq q^3$$
3.5. Construction of families with optimal cycle threshold via LPS graphs. We now present families of LPS graphs with optimal $CT(G)$. Let $p, q$ be primes satisfying $p, q = 1 \pmod{4}$. To construct a family of graphs with optimal threshold, our plan is to fix $p$ and consider an increasing sequence of primes $q_n$ with $(\frac{p}{q_n}) = -1$. For such a plan to work we need first to know that there is an infinite sequence of primes $q_n$ such that $q = 1 \pmod{4}$ and $(\frac{p}{q}) = -1$.

From quadratic reciprocity and our assumption that $p = 1 \pmod{4}$ it follows that $(\frac{p}{q}) = (\frac{p}{q})$. For any $x \in F_p$ such that $x \neq 0$, we have $x^2 = (-x)^2$ and $x \neq -x$. Therefore, there are precisely $(p-1)/2$ non zero quadratic residues and $(p-1)/2$ non residues modulo $p$. Pick some non residue $r$ and let $r'$ be such that $r' \equiv r \pmod{p}$ and $r' \equiv 1 \pmod{4}$. By the Chinese remainder theorem, such an $r'$ exists and is relatively prime (no common divisor) to $4p$. Any prime $q = r' \pmod{4p}$ will satisfy our requirements. It follows from Dirichlet’s theorem on primes in arithmetic progressions, that for each fixed prime $p > 2$ there are infinitely many primes $q$ with $q = 1 \pmod{4}$ and $(\frac{p}{q}) = -1$.

The proof that the cycle threshold for this family is optimal generalizes the girth lower bound argument for LPS graphs, given in [126].

**Theorem 46.** ([126] theorem 3.4) Let $p, q$ be primes as in the LPS construction. Assume that $(\frac{p}{q}) = -1$ and assume there exists a walk of length $k$ in $X_{p,q}$ corresponding to a normalized solution $x_0, x_1, x_2, x_3$ of (67), then the following hold.

1) All possible values of $x_0$ lie in the union of two arithmetic progressions with difference $2q^2$.
2) $k \geq 4 \log_p(q)$
3) $g(X_{p,q}) \geq (4/3) \log_p |V(X_{p,q})|$

**Proof** Consider a normalized solution of (67). Since

\[(70) \quad x_0^2 = p^k \pmod{q^2}\]

and $p$ is not a quadratic residue modulo $q$, $k$ must be even. The equation has only two solutions modulo $q$, say $x_0 = a, -a$. Consider the residue classes $a + jq$, $0 \leq j \leq q - 1$. We have $(a + jq)^2 \pmod{q^2} = a^2 + 2jq \pmod{q^2}$. Since $2$ is relatively prime to $q$, we see that modulo $q^2$ all the squares are different and they are a rearrangement of the classes $a^2 + kjq$, $0 \leq k \leq q - 1$. Since $p^k \equiv 1 \pmod{q^2}$ is of this form there is a unique element of the form $a + jq$ which solves equation (70). The same holds for $-a$, therefore there are $2$ solutions to (70), which in this case must be $p^{k/2} \pmod{q^2}$ and $-p^{k/2} \pmod{q^2}$. Combining this with the fact that $x_0$ is odd we obtain part 1. The solution $x_0 = p^{k/2}, x_1 = 0, x_2 = 0, x_3 = 0$ to equation (67) is not normalized, thus if $p^{k/2} \leq q^2$ there is no normalized solution with $x_0 = p^{k/2} \pmod{q^2}$, as the other possible (positive) values of $x_0 = p^{k/2} \pmod{q^2}$ are too large to satisfy (67). Therefore, we must have $x_0 = q^2 - p^{k/2}$ which
is not possible since \( x_0 \) must be odd. Hence, \( p^{k/2} > q^2 \), establishing part 2. Part 3 follows from part 2 and the inequality \( |V(X_{p,q})| \leq q^3 \) which follows from (69). \( q.e.d. \)

We need an estimate in terms of \( n \) on the number of Gaussian integers of a given norm.

**Theorem 47.** Let \( e(n) \) be the number of Gaussian integers of norm \( n \), i.e., the number of solutions to \( x^2 + y^2 = n \). Fix \( \varepsilon > 0 \), then there is a constant \( c > 0 \) such that for any \( n > n(\varepsilon) \) we have \( e(n) < cn^{\varepsilon} \).

**Proof:** We write \( n = \Pi n_i^{m_i} \). We have \( e(n) = \Pi e(n_i^{m_i}) \), so it is enough to show that for all but a finite number of pairs \( p_i, a_i \), we have \( e(p_i^{m_i}) < p_i^{m_i \varepsilon} \). If \( p_i^c > 2 \), then \( e(p_i^{m_i})/p_i^{m_i \varepsilon} = m_i + 1 / p_i^{m_i \varepsilon} < m_i + 1 / x_i^2 \leq 1 \) since \( 2^c \geq x + 1 \) for \( x \geq 1 \). For any fixed \( p \leq 2^{1/\varepsilon} \), and \( m \) large enough then \( p^{m / \varepsilon} > m + 1 \) as well and we are done. \( q.e.d. \)

The following result of Tillich and Zemor, [175], proves that the family of LPS graphs is optimal, see also [5].

**Theorem 48.** Fix a prime \( p = 1 \mod 4 \), put \( d = p + 1 \) and let \( q_n \) be an increasing sequence of primes all equal to 1 modulo 4 and satisfying \( (p, q_n) = 1 \). Let \( G \) be the family of graphs \( G_n = X_{p,q_n} \), then

\[
CT(G) = \frac{1}{d - 1}.
\]

**Proof:** Let \( c_{k,n} \) denote the number of cycles of length \( k \) in \( G_n \). Let \( r_{k,n} = r_{p,q_n}(k) \) be the number of normalized solutions to equation (67). Since the numbers of cycles of length \( k \) is bounded by the number of closed walks of length \( k \) we have

\[
c_{k,n} \leq r_{k,n} |V(G_n)| \leq r_{k,n} q_n^3.
\]

Any \( x_0 \) in a solution to equation (67) must satisfy \( |x_0^2| \leq p^k \) hence \( |x_0| \leq p^{k/2} \). By part 1 of Theorem 46 any such \( x_0 \) must belong to a union of two arithmetic progressions with difference \( 2q^2 \), hence there are at most \( p^{k/2}/q^2 + 2 \) possible values for \( x_0 \). By part 2 of Theorem 46 \( p^{m/2}/q^2 \geq 1 \), hence there is a constant \( A_1 \), say \( A_1 = 3 \), such that there are at most \( A_1 p^{k/2}/q_n^2 \) choices for \( x_0 \). Given one of these choices of \( x_0 \) we consider \( m = (p^k - x_0^2)/q_n = x_1^2 + x_3^2 \). Obviously \( |m| \leq p^k/q_n^2 \). We also have \( x_1^2 \leq m \). As \( x_1 \) can be either positive or negative but must be even, for each fixed \( x_0 \) there are at most \( \sqrt{m} \leq p^{k/2}/q_n \) possible choices for \( x_1 \). Having chosen \( x_1 \) as well we need to consider all solutions to \( x_2^2 + x_3^2 = m - x_1^2 \leq p^k \). By theorem 47 for any \( \varepsilon > 0 \) there is a constant \( A_\varepsilon \) such that there are at most \( A_\varepsilon p^{\varepsilon k} \) solutions to this equation. We thus have

\[
r_{p,q_n}(k) \leq A_1 (p^{k/2}/q_n^2)(p^{k/2}/q_n^2)^{-1}(A_\varepsilon p^{\varepsilon k}) \leq A_1 A_\varepsilon p^{(1+\varepsilon)k}/q_n^3
\]
solutions to equation (67). We deduce that \( c_{k,n} \leq A_1 A_\varepsilon (p^{(1+\varepsilon)^k}) \). Suppose, now, that the edges of the graph \( G_n \) are chosen randomly and independently with probability \( \rho \), forming the random subgraph \( G_n(\rho) \). Then the probability of a given cycle of length \( k \) to be chosen is \( \rho^k \). Let \( \rho = p^{-(1+\delta)} \) for some \( \delta > 0 \), and let \( \varepsilon \) be such that \( \delta > \varepsilon > 0 \). Let \( C_{n,\rho} \) be the random variable which counts the number of cycles in \( G_n(\rho) \) and \( D_{n,\rho} \) the random variable that indicates whether the graph \( G_n(\rho) \) has a cycle. Obviously \( C_{n,\rho} \geq D_{n,\rho} \), hence \( E(C_{n,\rho}) \) is an upper bound on the probability \( E(D_{n,\rho}) \) that a graph in the \( G_n(\rho) \) model contains a cycle. We have

\[
E(C_{n,\rho}) = \sum_{k=3}^{\infty} c_{k,n} \rho^k = \sum_{4 \log_\rho(q_n)}^{\infty} c_{k,n} \rho^k \\
\leq \sum_{4 \log_\rho(q_n)}^{\infty} A_1 A_\varepsilon (p^{(1+\varepsilon)})^k \\
\]

The last quantity tends to 0 since \( q_n \to \infty \) and by our choice of \( \varepsilon \). q.e.d.

The LPS construction produced families of \( d \)-regular graphs when \( d = p^{k+1} \) where \( p \) is a prime. In [137], M. Morgenstern generalized this construction to give families of \( p^k + 1 \)-regular graphs, where \( p \) is any prime and \( k \) is any positive integer, which again can be shown to have optimal cycle threshold.

4. Constructions for general degrees and graphs sizes

The families of graphs described in the previous section have asymptotically optimal cycle thresholds. However these families exist only if \( d = p^{k+1} \) where \( p \) is a prime. Moreover, if we denote the number of vertices of each graph in a family by \( n \), then even for degrees \( d \) as above the construction yields graphs only for a sparse set of values of \( n \). In our application, \( n \) is the number of disks and one would like to construct graphs whose subgraphs tend to be acyclic for all admissible values of \( d \) and \( n \).

Looking at the proof of Theorem 48 we notice that there are two properties which allowed us to prove optimality. The first is high girth and the second is that for any \( k \) there are at most \( O((d-1)^k) \) cycles of size \( k \) in the graph. Random \( d \)-regular graphs have the second property but not the first.

There are several ways to generate high-girth random or pseudo-random \( d \)-regular graphs of given size \( n \), see, for example, [34], [131], but it seems difficult to prove that they retain the second property.

Instead, we will consider, random \( d \)-regular graphs on \( n \) labeled vertices using the well studied configuration model, see [33], [101], and then modify them in order to eliminate the few short cycles they contain, keeping the property of having a relatively small number of longer cycles.
The generation of the random graph and the modifications will form a simple, efficient, randomized procedure to generate, for every fixed integer \( d \), and fixed real \( \epsilon > 0 \), and for any large integer \( n \) so that \( nd \) is even, a \( d \)-regular graph \( G \) on \( n \) vertices, such that a random subgraph of \( G \) obtained by keeping each edge of \( G \), randomly and independently, with probability \( \frac{1-\epsilon}{d-1} \), is acyclic with high probability.

The precise statement of the result is the following.

**Theorem 49.** Let \( d \geq 2 \) be a fixed integer, and let \( \epsilon > 0, \delta > 0 \) be positive reals. Define \( k_0 = k_0(\epsilon, \delta) = \lceil \frac{1}{\epsilon} \ln(\frac{1}{\delta}) \rceil \) and \( K = K(d, \epsilon, \delta) = 16d^{k_0} \). Then, for every integer \( n > K \) so that \( nd \) is even, there exists a \( d \)-regular graph \( G \) on \( n \) vertices, such that a random subgraph of \( G \) obtained by keeping each edge of \( G \), randomly and independently, with probability \( \rho = 1-\epsilon(d-1) \), is acyclic with probability at least \( 1-\delta - \frac{2K}{\epsilon}n^{-\epsilon/2} \log d \). Moreover, there is an efficient algorithm which uses random choices that produces a graph that has this property with probability at least \( 1/4 \). The probability of success can be made arbitrarily close to 1 by independent trials of the algorithm.

Before proving the result, we note both \( \epsilon \) and \( \delta \) in the theorem may be functions of \( n \). In particular, by taking them to be some functions that tend slowly to zero as \( n \) grows, e.g., \( \epsilon = \delta = 1/\log \log \log n \), we get a family of \( d \)-regular graphs with optimal cycle threshold \( d^2/d-2 \) for any \( d \).

**Proof:** Fix \( d, \epsilon, \delta \) and a sufficiently large \( n \) so that \( nd \) is even. Assume that \( d \geq 3 \), as the result for \( d < 3 \) is trivial. Let \( H = (V, E) \) be a random \( d \)-regular graph on \( n \) labeled vertices generated according to the configuration model described in [33], [101]. The graph is generated as follows. We consider the set \( W = \{v_{ij} : 1 \leq i \leq n, 1 \leq j \leq d \} \). We choose uniformly at random a pairing of all the elements of \( W \). This is also known as a perfect matching of \( W \). We then collapse each group \( V_i = \{v_{ij}, 1 \leq j \leq d \} \) to a vertex \( v_i \). The pairs of elements provide the edges of the graph. Notice that in the resulting graph, a vertex can have a loop, i.e., an edge onto itself. The number of configurations in this model, that is, the number of perfect matchings on \( W \), is \( (nd-1)!! = (nd-1)(nd-3)(nd-5) \cdots 1 \). For every integer \( k \) satisfying \( 1 \leq k \leq n \), the number of these configurations leading to a graph in which the vertices \( v_1, v_2, \ldots, v_k \) form a simple cycle of length \( k \) in this order is at most

\[
(d(d-1))^k(nd-2k-1)!!
\]

As there are \( \frac{1}{2k}n(n-1)(n-2) \cdots (n-k+1) \) potential simple cycles of length \( n \) on the vertices \( v_1, v_2, \ldots, v_n \), this implies that the expected number of simple cycles of length \( k \) in \( H \) is at most

\[
c(n, d, k) = \frac{1}{2k} \frac{n(n-1) \cdots (n-k+1)(d(d-1))^k}{(nd-1)(nd-3) \cdots (nd-2k+1)}.
\]
This number is at most \( \frac{1}{d^k} (d-1)^k \) for all \( k \geq 3 \), and at most \( \frac{1}{2k} (d-1)^k (1 + \frac{c}{n}) \) for some constant \( c > 0 \), when \( k = 1, 2 \).

Let \( C_k \) denote the number of cycles of length \( k \) in \( H \), and let \( \rho = \frac{1-\epsilon}{d-1} \) be as in the theorem. By linearity of expectation, the expected number of cycles of length at most \( k_0 \) in \( H \) is at most \( \sum_{k=1}^{k_0} c(n, d, k) < 2d^{k_0+1} = K/8 \), where \( K \) is as defined in the statement of the theorem. Therefore, with probability at least 7/8, the number of such cycles is at most \( K \).

Let \( k_1 = \frac{1}{2} \log_d n \). As in the previous paragraph, the expected number of cycles of length at most \( k_1 \) in \( H \) is at most \( \sum_{k=1}^{k_1} c(n, d, k) < 2d^{k_1+1} < \frac{1}{2} n^{2/3} \), hence, with probability at least 7/8, the number of such cycles is at most \( n^{2/3} \).

Let \( X = X(H) \) be the random variable defined as follows:

\[
X = \sum_{k > k_0} C_k \rho^k
\]

Again, linearity of expectation gives that the expectation of \( X \) is at most

\[
E(X) \leq \sum_{k > k_0} \frac{1}{2k} (d-1)^k \rho^k < \frac{1}{2k_0} \sum_{k > k_0} (1 - \epsilon)^k \leq \frac{\delta}{2}.
\]

Therefore, with probability at least 1/2, the value of \( X \) is at most \( \delta \). It follows that with probability at least 1/4, \( H \) has at most \( K \) cycles of length at most \( k_0 \), at most \( n^{2/3} \) cycles of length at most \( k_1 \), and the random variable \( X \) computed at \( H \) is at most \( \delta \). Fix such a graph \( H \). The desired graph \( G \) will be obtained from \( H \) by performing at most \( K \) switching operations, as described below, in order to destroy all cycles of length at most \( k_0 \), without creating any new cycles of length at most \( k_1 \), and without creating too many longer cycles.

Let \( e_1 = \{u_1, v_1\} \), \( e_2 = \{u_2, v_2\} \) be two edges in a graph \( H' \), where in each edge \( e_i \), \( u_i \) is considered the first vertex, and where \( e'_1 = \{u_1, u_2\} \), \( e'_2 = \{v_1, v_2\} \) are non-edges. The graph obtained from \( H' \) by switching \( e_1, e_2 \) is the graph \( H'' \) obtained from \( H' \) by deleting the edges \( e_1, e_2 \) and by adding the two new edges \( e'_1, e'_2 \). Note that if \( H' \) is \( d \)-regular, then so is \( H'' \). We claim that if the distance in \( H \) between the two edges \( e_1, e_2 \) is at least \( k_1 \), and \( e_2 \) lies in no cycle of length at most \( k_1 \), then the switching operation creates no new cycles of length at most \( k_1 \). Indeed, if a new cycle contains only one of the newly added edges, say \( \{u_1, u_2\} \), then it must contain a path in \( H' \) from \( u_1 \) to \( u_2 \), and by assumption, the length of any such path is at least \( k_1 \). If a new cycle contains both newly added edges, then it must contain either a path in \( H' \) from \( u_1 \) to \( v_2 \), or a path in \( H' \) from \( u_2 \) to \( v_1 \), and in both cases the resulting cycle is of length exceeding \( k_1 \). Another simple observation is the fact that for any \( k \) and for any edge \( e \) in a \( d \)-regular graph, \( e \) can lie in at most \( (d-1)^k \) cycles of length \( k \), as the number of walks of length \( k-1 \) starting at a vertex is bounded by \( (d-1)^{k-1} \). Since any switching operation
adds two new edges, it can add at most $2(d-1)^k$ new cycles of length $k$, for any $k$.

Returning to our graph $H$, we now modify it to obtain the desired graph $G$ as follows. Starting with $H$, as long as our graph contains a cycle of length at most $k_0$, pick an arbitrary edge $e_1$ in it, and pick another edge $e_2$ of distance at least $k_1$ from $e_1$ which does not lie on a cycle of length at most $k_1$. Then switch $e_1, e_2$. As this process creates no new cycles of length at most $k_1$, throughout the process our graph contains at most $n^2/3$ cycles of length at most $k_1$. Therefore, at most $n^2/3 k_1 < n^{2/3} \log n$ edges lie on such cycles, and as the number of edges within distance $k_1$ from $e_1$ is at most $2d^{k_1} = 2\sqrt{n}$, there is always a valid choice for $e_2$. Each such switching operation destroys the cycle of length at most $k_0$ through $e_1$, and hence this process must terminate after at most $K$ steps. By the discussion above, this gives a graph $G$ of girth exceeding $k_0$, in which the number of cycles of each length $k \leq k_1$ is precisely $C_k$, the number of cycles of that length in $H$. Moreover, the number of cycles of length $k$ in $G$ for larger values of $k$ is at most $C_k + 2K(d-1)^k$.

Suppose, now, that $G(\rho)$ is a random subgraph of $G$ obtained by picking each edge of $G$, randomly and independently, with probability $\rho$. Then the expected number of simple cycles in $G(\rho)$ is at most

$$\sum_{k=k_0+1}^{k_1} C_k \rho^k + \sum_{k>k_1} (C_k + 2K(d-1)^k) \rho^k \leq \delta + \sum_{k>k_1} 2K(d-1)^k \rho^k$$

$$= \delta + 2K \sum_{k>k_1} (1-\epsilon)^k \leq \delta + \frac{2K}{\epsilon} n^{-\epsilon/2 \log d}.$$ 

It follows that the probability that $G(\rho)$ contains a cycle does not exceed $\delta + \frac{2K}{\epsilon} n^{-\epsilon/2 \log d}$, as needed.

The randomized algorithm to generate $G$ is simple: generate $H$, find all its cycles of length at most $k_1$ (by checking all walks of that length), and then perform the switchings as in the proof. This completes the proof. \textit{q.e.d.}

4.1. Some Experimental results. We have chosen to test the ability of various configuration graphs to support active sets for three different parameter settings of $d$ and $n$ in which we have good explicit candidates for large girth graphs. Based on the results of the previous section we expect the large girth graphs to do well. In all three cases we compared our explicitly constructed configuration to a set of 800 randomly chosen configurations with the same parameters. Each graph $G$ was tested as follows. We considered 250 random subgraphs of $G$ with $k$ edges, for a given $k$. We then counted how many of these subgraphs had only tree components. Dividing by the number of experiments, we obtained empirical estimates on the probability $p_k$ of having only tree components. We varied $k$ in reasonable small increments in the range where $p_k$ is expected to pass from a value close
to 1 to a value close to 0. We also plotted the probabilities $p_k$ for random subgraphs of the complete graphs on $n$ vertices, that is random graphs with $n$ vertices and $k$ edges.

The results are presented in the following tables.

The first column ($k$) presents the value of $k$.

The second column (Special) represents the values of $p_k$ obtained by the specially designed graph.

The third column (AverageRand) represents the average value of $p_k$ obtained from the 800 randomly chosen $d$ regular graphs with $n$ vertices.

The fourth column (MaxRand) represents the maximal value of $p_k$ among the 800 values obtained from the randomly chosen graphs.

The last column (Complete) presents $p_k$ for the complete graph with $N$ vertices.

### 4.1.1. $d=3$, $n=70$

The special high girth graph we consider in this case is described in [23]. It has girth 10 which, as shown in [140], is the maximal value attainable by 3 regular graphs of this size. We refer the reader to [29] for much more information on constructing 3 regular graphs with high girth.

<table>
<thead>
<tr>
<th>$k$</th>
<th>Special</th>
<th>AveRand</th>
<th>MaxRand</th>
<th>Comp</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>0.98</td>
<td>0.76</td>
<td>0.91</td>
<td>0.26</td>
</tr>
<tr>
<td>46</td>
<td>0.96</td>
<td>0.71</td>
<td>0.90</td>
<td>0.21</td>
</tr>
<tr>
<td>48</td>
<td>0.96</td>
<td>0.65</td>
<td>0.84</td>
<td>0.15</td>
</tr>
<tr>
<td>50</td>
<td>0.89</td>
<td>0.58</td>
<td>0.77</td>
<td>0.05</td>
</tr>
<tr>
<td>52</td>
<td>0.76</td>
<td>0.50</td>
<td>0.74</td>
<td>0.03</td>
</tr>
<tr>
<td>54</td>
<td>0.72</td>
<td>0.42</td>
<td>0.64</td>
<td>0.00</td>
</tr>
<tr>
<td>56</td>
<td>0.58</td>
<td>0.33</td>
<td>0.55</td>
<td>0.00</td>
</tr>
<tr>
<td>58</td>
<td>0.49</td>
<td>0.23</td>
<td>0.40</td>
<td>0.00</td>
</tr>
<tr>
<td>60</td>
<td>0.31</td>
<td>0.15</td>
<td>0.31</td>
<td>0.00</td>
</tr>
<tr>
<td>62</td>
<td>0.16</td>
<td>0.08</td>
<td>0.18</td>
<td>0.00</td>
</tr>
<tr>
<td>64</td>
<td>0.10</td>
<td>0.03</td>
<td>0.08</td>
<td>0.00</td>
</tr>
</tbody>
</table>

### 4.1.2. $d=3$, $n=126$

In this case our special graph is a graph of girth 12 whose construction appears in [28]. This graph has exceptionally large girth for its size and is the unique graph of such girth with 126 vertices.

<table>
<thead>
<tr>
<th>$k$</th>
<th>Special</th>
<th>AveRand</th>
<th>MaxRand</th>
<th>Comp</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>0.98</td>
<td>0.74</td>
<td>0.93</td>
<td>0.16</td>
</tr>
<tr>
<td>84</td>
<td>0.93</td>
<td>0.67</td>
<td>0.89</td>
<td>0.09</td>
</tr>
<tr>
<td>88</td>
<td>0.91</td>
<td>0.59</td>
<td>0.82</td>
<td>0.04</td>
</tr>
<tr>
<td>92</td>
<td>0.88</td>
<td>0.51</td>
<td>0.74</td>
<td>0.01</td>
</tr>
<tr>
<td>96</td>
<td>0.72</td>
<td>0.40</td>
<td>0.60</td>
<td>0.01</td>
</tr>
<tr>
<td>100</td>
<td>0.61</td>
<td>0.29</td>
<td>0.47</td>
<td>0.00</td>
</tr>
<tr>
<td>104</td>
<td>0.38</td>
<td>0.18</td>
<td>0.31</td>
<td>0.00</td>
</tr>
<tr>
<td>108</td>
<td>0.22</td>
<td>0.09</td>
<td>0.19</td>
<td>0.00</td>
</tr>
<tr>
<td>112</td>
<td>0.08</td>
<td>0.03</td>
<td>0.07</td>
<td>0.00</td>
</tr>
</tbody>
</table>
4.1.3. $d=6$, $n=2184$. Our special graph is the LPS graph with $p=5$ and $q=13$.

\[
\begin{array}{|c|c|c|c|c|}
\hline
k & Special & AveRand & MaxRand & Comp \\
\hline
950 & 0.98 & 0.85 & 0.91 & 0.65 \\
1000 & 0.98 & 0.82 & 0.89 & 0.59 \\
1050 & 0.95 & 0.77 & 0.86 & 0.45 \\
1100 & 0.94 & 0.73 & 0.82 & 0.36 \\
1150 & 0.86 & 0.66 & 0.76 & 0.22 \\
1200 & 0.82 & 0.58 & 0.67 & 0.09 \\
1250 & 0.68 & 0.48 & 0.57 & 0.05 \\
1300 & 0.53 & 0.36 & 0.48 & 0.00 \\
1350 & 0.37 & 0.23 & 0.30 & 0.00 \\
1400 & 0.20 & 0.11 & 0.18 & 0.00 \\
1450 & 0.07 & 0.03 & 0.06 & 0.00 \\
1500 & 0.02 & 0.01 & 0.03 & 0.00 \\
\hline
\end{array}
\]

Notes for chapter 4

(1) In the chapter we discussed, the case where there are are $d$ files on each disk. Since the capacity of disk drives is very large, we are implicitly considering very large files. The number of files that can be supported is obviously at most $n$, the number of disks. If we consider individual requests instead of files, then the size of a request is usually small and there is a very large number of data pieces which can be requested on each disk. From this point of view, where requests are managed individually and not as part of a larger entity such as a file or LUN, there is a huge number of possible data locations on a disk and in terms of modeling we should abandon the concept of a configuration graph. Instead, we simply assume that each of $m$ outstanding requests can be serviced from either of $h$ randomly chosen disks, those which contain a copy of the requested data. The problem becomes a load balancing problem of assigning each request to a single disk, so that the maximum number of requests $k$ which are assigned to a given disk is minimized. This model is particularly relevant if we replace disk drives by a RAM (random access memory) storage device where there is no mechanical motion involved and so the service time of requests to a device is proportional to their number. The case we considered in this chapter corresponds to $h=2$ (two disks containing each piece of data) and $k=1$, at most one entity (file or request) are assigned to each disk. In this case as the degree $d$ goes to $\infty$ we see that the number of supported requests tends to $n/2$ and indeed this is the threshold in Erdos-Renyi theory where a complicated component containing many cycles first appears. More generally, one assumes
that there are $m = cn$ requests and one looks for a threshold $c_{k,h}$ such that for $c < c_{k,h}$, w.h.p., as $n \to \infty$ we can assign $m = cn$ requests, each residing on $h$ i.i.d. uniformly chosen disks, to $n$ disks, so that each disk is assigned at most $k$ requests, while, if $c > c_{k,h}$ it is not possible to make such an assignment. In this language we have $c(1,2) = 1/2$.

Obviously $c(k,h) \leq k$. In [154] it was shown that $c(k,2) \geq k - 1$.

The analysis of these types of models shares resemblance with our upper bound on the threshold $\rho$, but is more complicated. To begin with, for $h > 2$ we model the problem by a hyper-graph rather than a graph. An $h$ hyper-graph $U$ on a set $V[U]$ of $n$ vertices is a set $H(U)$ of hyper-edges which are subset of $U$ of size $h$, the case $h = 2$ corresponding to edges in a graph. The degree of a vertex is the number of hyper-edges it is contained in. The $l$-core of a hyper-graph consists of a maximal subset of vertices such that each vertex in the subset is in at least $l$ hyper-edges whose vertices are also in the subset. The hyper-edge density of a hyper-graph is the ratio of the number of hyper-edges to vertices. If w.h.p. the hyper-edge density of the $k+1$ core will be more than $k$, then we will have to assign to one of the vertices (storage devices) in the core more than $k$ hyper-edges (requests), hence w.h.p., the $m$ requests will not be supported. We conclude that the threshold $c(k,h)$ will be at most the threshold $\tilde{c}(k,h)$ for having hyper-edge density $k$ in the $k+1$ core. It is shown in [64] that for $k = 1$ we have $c(1,h) = \tilde{c}(1,h)$ and equality is conjectured to hold for all $k$. Results for $k > 1$ can be found in [47]. The value of $\tilde{c}(k,h)$ can be characterized (and computed numerically) by relating the expected number of vertices and hyper-edges in cores to a related branching process.

(2) The LPS graphs, at the time of their introduction, held the world record for large girth. However, the main interest in these graphs was that they had another extremal property that we now explain.

**Definition 90.** Given a graph $G$ with $n$ vertices, we define the (edge) expansion ratio $h(G)$ of $G$ to be

$$\min_{A \subseteq V(G), |A| \leq n/2} \frac{|E(A,V(G) - A)|}{|A|}$$

We say that a family $G_i$ of $d$ regular graphs with $|V(G_i)| \to \infty$ is said to be a family of (edge) expander graphs if there exists some $\varepsilon > 0$ such that $h(G_i) \geq \varepsilon$ for all $i$.

This definition first appeared implicitly in [116] and explicitly in [143], where expander graphs were related to the construction of efficient communication networks. It is not difficult to show by
probabilistic arguments that a family of randomly chose \( d \) regular graphs will be an expander, however it is difficult to make explicit constructions.

Expander graphs have numerous important applications in theoretical computer science, see [97], and more recently in pure mathematics, [125]. In the computer science applications it is important that the graphs be explicitly computable, since these graphs are used to produce efficient algorithms and network designs. The mathematical applications also involve non-random constructions involving groups. The first explicit construction was made by Margulis, [129]. The proof of the expander property was based on representation theory.

Let \( A \) be the adjacency matrix of a \( d \) regular graph \( G \). Since \( A \) is symmetric, all its eigenvalues are real. Since the graph is \( d \) regular, the matrix \( A/d \) is doubly stochastic and hence for any eigenvalue \( \lambda \) of \( A \) we have \(|\lambda| \leq d \). It is easy to see that \( d \) is always an eigenvalue since the vector \((1, 1, \ldots, 1)\) is an eigenvector for this eigenvalue. Let \( \lambda_1 \) be the second largest eigenvalue. The following result relates the expansion of a graph to the value of \( \lambda_1 \), see [56] or [97] for a proof.

**Theorem 50.** A connected graph \( G \) satisfies

\[
\sqrt{(d + \lambda_1)(d - \lambda_1)} \geq h(G) \geq (d - \lambda_1)/2
\]

The first inequality is due to Dodziuk, [66], while the second is due to Alon and Milman, [7]. By the theorem, constructing a family of expander graphs is the same as constructing a family with \( \lambda_1(G) \) being bounded away from \( d \), i.e., \( \lambda_1(G_n) \leq d - \delta \) for some \( \delta > 0 \). The smaller \( \lambda_1 \) is, the better the expansion guarantee.

Following the theorem, we shift our attention to families with small \( \lambda_1 \). The following theorem, due to Alon and Boppana, places a limit on the smallness of \( \lambda_1 \), see [56] for a proof.

**Theorem 51.** Let \( G_n \) be a family of \( d \) regular graphs with \( V(G_n) \to \infty \). Let \( \lambda(G_n) \) be the largest absolute value of an eigenvalue of the adjacency matrix of \( G_n \), different from \( d, -d \). Then, for any \( \varepsilon > 0 \) and \( n \) large enough we have

\[
\lambda(G_n) \geq 2\sqrt{d - 1} - \varepsilon
\]

Given the theorem we have the following definition which characterizes the extreme families in terms of smallness of \( \lambda_1 \).

**Definition 91.** We say that a graph \( G \) is a **Ramanujan graph** if \( \lambda(G) \leq 2\sqrt{d - 1} \).

The following result is considered the main theorem of [126], a substantial portion was also proved independently in [130].
Theorem 52. The $LPS$ graphs are Ramanujan.

The $LPS$ graphs were the first family of Ramanujan graphs to be shown to exist. There are no random graph arguments that show the existence of families of Ramanujan graphs. All known families of Ramanujan graphs come from number theoretic or representation theoretic arguments. Families of Ramanujan graphs are not known to exist when $d$ is not of the form $p^k + 1$, $p$ being prime.
CHAPTER 5

On queues and numbers

In this chapter, the last of the book, we consider performance analysis, more generally, outside the scope of storage systems.

The area of mathematics which is most strongly associated with performance analysis is queueing theory. One of the things we hate doing is waiting. We wait in line at the bank or supermarket, we may have to wait for the computer to do some job that we asked it to do, because it has some other tasks. We may wait for a file to download and the reader can come up with many other scenarios. In general, we want to analyze systems with "customers", which arrive into the system at some given times. Each customer requires some "job" to be done. The jobs are handled by "servers" which require a certain amount of service time to handle the job. Sometimes customers arrive into the system and find out that no one is available to service them, or nobody wants to service them. Furthermore, in some cases service may be temporarily disrupted and resumed later. Whenever the customer is not being serviced we may say that the customer is waiting. Whenever that happens, queues of waiting customers start to accumulate in the system and one has to manage them somehow, that is, to decide which customers should wait and which should be serviced, given the system (server) resources. This is often done with some target in mind, usually, trying to minimize some variant of the average waiting time, or completion time. To study such systems mathematically we need to have information on the nature of the arrival times and job sizes of the customers. One method is to specify for each customer, when it arrived and what were their job requirements. This is the analogue of a trace. Another method is to model the arrival times and job requirements using some stochastic model as we have done with the IRM model in chapter 1. The theory of scheduling uses traces as input and attempts to find queue management schemes that will have good performance for any conceivable input trace. Queueing theory on the other hand, models the input data using stochastic processes and tries to analyze the resulting system as accurately as possible under the specific model assumptions. One can say that scheduling theory provides an analysis which is good under all circumstances but less accurate, while queueing theory can provide (in some cases) very detailed information, but that information is only as good as the modeling of the data. Both theories have very nice results, but as we have done in the rest of the book we will
consider the stochastic approach, namely queueing theory. For a taste of scheduling, we refer the reader to the survey paper [147].

There are many excellent introductory texts to queueing theory. The two volume book of Kleinrock, [113] is especially recommended. Consequently, instead of giving an overview of the basics of queueing theory we will consider a particular problem and introduce the subject through a very specific result. Analyzing the problem and putting it in context will lead us to consider some basic results in queueing theory but will also lead us in parallel to consider some basic results in number theory, hence, the title of the chapter. While the problem we will study has been considered mostly in the context of computer servers, it will be convenient to explain it in the context of managing a mini-market with two checkout counters, one of them acting as an express line.

1. Managing a mini-market

We consider a mini-market with two checkout counters, numbered 1 and 2.

1.1. The ingredients of the mini-market queueing system. In order to mathematically analyze a queueing system we need some information that we now specify.

**Definition 92.** We say that a queueing system has Poisson arrivals with rate $\lambda$ if customers may arrive after some time $T = 0$ and the number of customers arriving at any given time interval $I = [a, b]$ has a Poisson distribution with parameter $\lambda(b - a)$, with arrivals in disjoint time intervals being independent.

1.1.1. Server strength assumptions. Our mini-market has two checkout counters where service is provided. In the queueing theory and scheduling literature, service areas or service providers are termed machines, or servers, or hosts. We will use the latter two terms.

**Definition 93.** Consider a queueing system with $h$ servers, numbered $1, ..., h$.

We say that the servers are coupled if we can assign to any job a number $t > 0$, called the job size and there exist functions $T_i(t)$ for $1 \leq i \leq h$ and $t \geq 0$ such that any job of size $t$ takes $T_i(t)$ time to process on server $i$.

We say that the servers of the system, are linearly coupled if we can assign to server $i > 1$ a server speed $s_i$ such that if the processing time of a job on server 1 is $t$ then the service time of the same job on server $i > 1$ is $t/s_i$. We assume that $T_1(t) = t$, namely, that the job size is the processing time on server 1 and we define for consistency $s_1 = 1$.

We say that a queueing system has identical servers if a job that takes $t$ time to process on one server also takes $t$ time to process on any other server. In other words, if all speeds are equal to 1.
Definition 94. Assume that in a queueing system with coupled hosts, the job-size of the i’th customer is a random variable $X_i$ which is independent of any other random variable which defines the queueing process and that all $X_i$ have a common distribution $X(t) = Pr(X_i \geq t)$. Under these assumptions, the distribution $X$ is known as the job-size distribution.

In the mini-market situation we will assume a linearly coupled model. In more detail, there are two workers in the mini-market Alice, to be denoted by the letter $A$ and Bob, or $B$ for short, with possibly different speeds. Alice and Bob are working at the checkout counters. As in the definition of linear coupling we assume without loss of generality that the speed of Alice is fixed to be 1 and the speed of Bob will be denoted by $b$. We observe that the parameter $b$ is equivalent to the parameter $1/b$ if we switch the names Alice and Bob, hence we may assume without loss of generality that $b \leq 1$, or in other words that Alice is the faster worker.

To define an express line we will need a cutoff $s$. We will send all customers whose checkout processing time $T$ satisfies $T \leq s$ to checkout counter 1 while those with processing time $T > s$ will be sent to checkout counter 2. It may be argued that we usually do not know the processing time $T$ of a customer, but to simplify matters we will assume that we do know.

The cutoff $s$ will always be with respect to the processing time of Alice. This is consistent with the normalization of her speed to be 1.

Since Alice and Bob are not equally capable in general we have to decide who serves customers in the express checkout counter which is counter number 1. We will describe this decision by a permutation $\sigma$ on two elements. The arrangement whereby Alice serves customers in the express counter and Bob is assigned to the other counter will be assigned to the identity permutation, while the other assignment in which Alice and Bob switch counters will correspond to the non-identity permutation. If $\sigma$ is a permutation on two elements, then we denote by $\bar{\sigma}$ the other permutation. The parameters $\lambda, X, s, b, \sigma$, specify completely a mini-market queueing system. However, it is more convenient to replace the customer arrival rate $\lambda$ by another parameter which takes into account the processing power of the servers.

Definition 95. The utilization $\rho$ of a queueing system with a single server, arrival rate $\lambda$ and job size distribution $X$ is defined as $\rho = \lambda E(X)$. For the mini-market system described above we define the utilization as $\rho = \frac{\lambda E(X)}{1+b}$.

In a single server system the utilization tells us the portion of time in which the system is busy serving customers. We recall that we are assuming an arrival rate of $\lambda$ customers per time unit on average. Since Alice has speed 1, the average size of a customer’s processing time, w.r.t. Alice, is $E(X)$, the average of the job-size distribution $X$. Similarly, on average a job takes Bob $E(X)/b$ time to process. This means that Alice can process on average $1/E(X)$ customers per time unit, while Bob can handle on average
5. ON QUEUES AND NUMBERS

Since they are working in parallel, they can handle up to \((1 + b)E(X)\) customers. Since they are working in parallel, they can handle up to \((1 + b)E(X)\) customers on average per time unit. If the customer arrival rate \(\lambda\) is larger than the average service rate \((1 + b)E(X)\), then regardless of the cutoff value, the queue at one of the counters will explode in size leading to an unstable system with infinitely large average waiting time. If on the other hand the ratio of customer arrival rate to customer service rate \(\frac{\lambda E(X)}{1+b}\) satisfies \(\rho < 1\), then, as the analysis will show, we can find a cutoff \(s\) for which the queueing systems will stabilize as a stochastic process.

1.2. The mini-market management problem. We will state three management problems for a mini-market queue with an express counter.

**Problem 1**: Assume that Alice and Bob are equally capable workers, i.e., \(b = 1\). In this case the permutation \(\sigma\) is immaterial. Let \(E(W)(X, \rho, s)\) be the average waiting time of customers in the mini-market queueing system (customer service time at the counter not included).

Given a pair \(X, \rho\), define \(E(W)_{opt} = \min_s E(W)(X, \rho, s)\). It will be shown later that the minimum actually exists. We let \(s_{opt} = s_{opt}(X, \rho)\) be an optimal cutoff value, namely, \(E(W)(X, \rho, s_{opt}) = E(W)_{opt}\). We will show later on that if the job size distribution is strictly increasing, then, there is only one optimal value.

The problem is to compute the value of the optimal cutoff \(s_{opt}(X, \rho)\).

**Problem 2**: We assume that \(b \neq 1\), i.e., Alice and Bob are not equally capable. Fixing \(X, \rho, b\), let \(E(W)_{opt}(\sigma) = \min_s E(W)(X, \rho, b, \sigma, s)\).

The question is whether \(E(W)_{opt}(\sigma) = E(W)_{opt}(\bar{\sigma})\). Stated otherwise, given \(X, \rho, b\), does it matter whether Alice works at the express counter or Bob?

In case it matters we can also ask which assignment should be preferred, should we place the faster worker, Alice, at the express (small jobs) counter or the other way around?

**Problem 3**: Continuing the scenario of problem 2, for fixed parameters \(X, \rho, b\), define \(s_{opt}(\sigma)\) to be a cutoff for which \(E(W)_{opt}(X, \rho, b, \sigma, s) = E(W)(X, \rho, b, \sigma, s_{opt})\). Suppose we know \(s_{opt}(\sigma)\), can we easily compute \(s_{opt}(\bar{\sigma})\) from it? In other words, suppose Alice has been working at the express counter for a while and that during this time the mini-market manager has been able to determine \(s_{opt}\) via trial and error or some other method. Suppose that one day Alice and Bob get bored and decide to switch counters, can the manager compute the new optimal cutoff from the old one?

After all this build up, we are sorry to report that for general parameters \(X, \rho, b\), the answers to these problems are tedious or take a negative form. Take for example \(X\) to be an exponential distribution, \(Pr(X \geq t) = e^{-\mu t}\), with some parameter \(\mu\) which satisfies \(\rho = \frac{\lambda(1+b)}{\mu} < 1\). For many queueing
1. MANAGING A MINI-MARKET

systems this particular choice of job-size distribution is easiest to analyze and leads to the cleanest results. In the mini-market case, the answer to problem 1 can only be determined numerically and depends on $\rho$. The answer to problem 2 is that it does matter who works at the express counter. For example, numerically, it has been determined that when $\rho = 0.8$ and $b = 1/2$ it is better to assign the slower worker, Bob, to the express counter, [73].

For other job-size distributions, the opposite is true and the faster worker should handle the express counter. In addition, it is not clear that the result does not depend on $\rho$ and $b$.

Regarding problem 3, there does not seem to be a nice formula relating an optimal cutoff to the optimal cutoff after the workers have switched counters.

1.3. A positive result. While the general situation seems bleak, we are happy to report that for some job-size distributions the answer to all three problems is simple and does not depend on the values of $\rho$ and $b$. A mini-market manager who is lucky enough to observe one of these special job-size distributions will have a very easy time managing the mini-market.

To illustrate this, we introduce a family $X_{A,B}$ of job-size distributions, parametrized by a pair of integers $A, B \in \mathbb{Z}$. The integers $A, B$ must satisfy some mild conditions. They must satisfy the inequality

$$\Delta_{A,B} = 4A^3 + 27B^2 \neq 0$$

In addition, for any prime $p$, if $p^4$ divides $A$ then $p^6$ does not divide $B$.

Using $A$ and $B$ we define a sequence of integers $a_n, n = 1, 2, ...$. We let $a_1 = 1$. Given $n, m$ with no common divisor we let $a_{nm} = a_n a_m$. By the unique decomposition of positive integers into a product of prime numbers we are left with the task of defining $a_{p^k}$ for all prime powers $p^k$. Consider the equation

$$(72) \quad y^2 = x^3 + Ax + B$$

Let $N_{p^k}$ denote the number of solutions to the equation (72), where we think of the integers $A, B$ as numbers modulo $p$ and $x, y$ are in the field with $p^k$ elements. Define $a_{p^k} = p^k - N_{p^k}$. Let $q = q(t) = e^{-2\pi t}$ and define the power series $g_{A,B} = \sum_{n=1}^{\infty} a_n q^n$. Let $\tilde{f}_{A,B}(t) = g_{A,B}^2(t)$.

**Theorem 53.** The following statements hold

1) The integral $\int_0^\infty \tilde{f}_{A,B} dt$ exists, thus we can define a density function $f_{A,B}(t) = \tilde{f}_{A,B}(t) / (\int_0^\infty \tilde{f}_{A,B} dt)$.

Consider the mini-market queueing system with job-size density function $f_{A,B}(t)$, and parameters $\rho, b$, then,
1) Let $b = 1$. There is an explicitly computable integer $N_{A,B}$ which is divisible by the same primes as $\Delta_{A,B}$ and such that the optimal cutoff is given by $s_{opt} = 1/\sqrt{N_{A,B}}$, independent of $\rho$.

2) For any $\rho$ and $b$ it does not matter which assignment of workers to counters is chosen, namely,

$$E(W)_{opt}(f_{A,B}(t), \rho, b, \sigma) = E(W)_{opt}(f_{A,B}(t), \rho, b, \bar{\sigma})$$

3) The following formula relates the optimal cutoffs with respect to $\sigma$ and $\bar{\sigma}$,

$$s_{opt}(f_{A,B}(t), \rho, b, \sigma) = \frac{1}{N_{A,B}s_{opt}(f_{A,B}(t), \rho, b, \bar{\sigma})}$$

independently of $\rho$ and $b$.

As can be observed from the theorem, managing the mini-market when $f_{A,B}(t)$ is the job-size density function could not be easier. The theorem is proved by putting together several results from queueing theory with several results from number theory. Most of the chapter, is devoted to explaining this theorem in its natural queueing theoretic and number theoretic contexts and how the two end up mixing together. At the end we consider a few more related facts about mini-market like queues.

1.4. A family of involutions. The central figure in our story is a family of involutions on functions of $t > 0$.

**Definition 96.** Let $f(t)$ be a function on $t > 0$. We define the Hecke involution with parameters $k, N$ by the formula

$$H_{k,N}(f)(t) = \hat{f}_{k,N}(t) = N^{-k/2}t^{-k}f(1/Nt)$$

The parameter $k$ will be called the **weight** of the involution $H_{k,N}$, while the parameter $N$ will be called the **level**.

If $I$ is the interval $(a, b), [a, b), (a, b], [a, b] \subset (0, \infty)$ then we define the **level $N$ dual of $I$**, denoted $I_N$, to be the image of $I$ under the involution $t \rightarrow 1/Nt$.

It is easy to check that $H_{k,N}$ is indeed an involution, i.e., when applied twice we obtain the identity map on functions. Obviously for any $N, I \rightarrow I_N$ is an involution on intervals.

The involutions $H_{k,N}$ have historically played an important role in number theory, going back to the work of Jacobi, Riemann and many others. We will show that they are also relevant to queueing theory. We begin their story on the number theoretic side.
1.5. Analytic continuation and functional equation for the Riemann zeta function.

**Definition 97.** The Riemann zeta function is given by the series

\[ \zeta(s) = \sum_n n^{-s} \]

which converges for complex values of \( s \) which satisfy \( \text{Re}(s) > 1 \).

The (Riemann) zeta function was first considered by Euler, [70]. Euler proved his famous product formula

\[ \zeta(s) = \prod_p \frac{1}{1 - p^{-s}} \]

where the product is taken over all primes.

The product formula already suggests that there is a relation between the properties of the function \( \zeta(s) \) and the distribution of prime numbers. The \( \zeta \) function was considered in depth in Riemann’s seminal paper, [150], hence its name. Since Riemann’s work, the properties of primes are studied in many cases through an analysis of the zeta function.

The first order of business was to provide a natural (complex analytic) extension of the definition of the function from its original domain of convergence \( \text{Re}(s) > 1 \) to an analytic function (having Taylor series expansions at all points) on the entire complex plane except for \( s = 0 \) and \( s = 1 \). In addition, Riemann proved a functional equation relating the values of \( \zeta(s) \) and \( \zeta(1 - s) \).

1.5.1. The functional equation of the theta function. One of the methods which Riemann gave for achieving both goals (analytic continuation and functional equation) was to relate the zeta function to another function with a different functional equation, Jacobi’s theta function, which we considered in chapter 4. We recall the definition of the theta function. Let \( q = e^{-\pi t} \).

We define

\[ \theta(t) = \sum_{n=-\infty}^{\infty} q^{n^2} = 1 + 2 \sum_{n \geq 1} q^{n^2} \]

Since \( q < 1 \) when \( t > 0 \), the series obviously converges for \( t > 0 \).

Jacobi proved that the \( \theta \) function satisfies the following functional equation.

**Theorem 54.** The theta function satisfies

\[ \theta(t) = t^{-1/2} \theta \left( \frac{1}{t} \right) \]

equivalently \( H_{1/2,1}(\theta) = \theta \).

**Proof:** For a given \( t \), let

\[ f_t(x) = e^{-\pi tx^2} = e^{-\pi(\sqrt{t}x)^2} \]
By definition

\[ \tilde{f}(y) = \int_{\mathbb{R}} e^{-\pi (\sqrt{t}x)^2} e^{-2\pi iyx} \, dx \]

Making the change of variable \( \tilde{x} = \sqrt{t}x \), and using the invariance of \( e^{-\pi x^2} \) under the Fourier transform, see the appendix, we obtain,

\[ \tilde{f}(y) = \int_{\mathbb{R}} e^{-\pi \tilde{x}^2} e^{-2\pi i \frac{y\tilde{x}}{\sqrt{t}}} \frac{d\tilde{x}}{\sqrt{t}} = t^{-1/2} e^{-\pi (y/\sqrt{t})^2} = t^{-1/2} e^{-\pi y^2/t} \]

Applying the Poisson summation formula, see the appendix, to \( f_t(x) \) yields Jacobi’s functional equation. \( q.e.d. \)

1.5.2. The Mellin transform. We want to relate Jacobi’s theta function to Riemann’s zeta function. The relation will be through the notion of a Mellin transform.

**Definition 98.** Let \( f(t) \) be a complex valued function defined on the non-negative reals \((0, \infty)\). We define for any complex number \( s \), the Mellin transform of \( f \) evaluated at \( s \) by the formula

\[ L_f(s) = \int_0^\infty f(t) t^s \frac{dt}{t} \]

whenever it exists. Given an interval \( I \subset (0, \infty) \) we also define the incomplete Mellin transform with respect to \( I \) by the formula

\[ L_{f,I}(s) = \int_I f(t) t^s \frac{dt}{t} \]

The following basic computation (observation) of Riemann relates the incomplete Mellin transforms of \( f \) and of \( \hat{f}, k,N \). While it is trivial, it is of central importance for the rest of the chapter.

**Lemma 16.** Let \( k, N \) be some weight and level. Let \( f \) be a function on the positive reals and let \( I \subset (0, \infty) \) be some interval, then, \( L_{\hat{f}, I}(s) \) exists if and only if \( L_{f, I}(k - s) \) exists and we have

\[ L_{\hat{f}, I}(s) = N^{k/2-s} L_{f, I}(k - s) \]

**Proof:** Letting \( \tau = 1/\sqrt{N}t \), we have \( d\tau = -\frac{dt}{t} \). Then,

\[ L_{f,I} = \int_I \hat{f}(t) t^s \frac{dt}{t} = \int_I N^{-k/2} t^{-k} f(1/\sqrt{N}t) t^s \frac{dt}{t} \]

\[ = N^{-k/2} \int_I \left( \frac{1}{\sqrt{N}} \right)^s \frac{f(\tau)}{\tau^k} \frac{d\tau}{\tau} = N^{k/2-s} L_{f,I}(s - k) \]

as required. \( q.e.d. \)
1. MANAGING A MINI-MARKET

1.5.3. Relating the theta and zeta functions through the Mellin transform. We want to claim that the Mellin transform of the theta function is closely related to the Riemann zeta function, however, the Mellin transform of the theta function does not converge. The problem is that the theta function has a constant term, hence, \( \lim_{t \to \infty} \theta(t) = 1 \), and so, for \( s > 0 \) we will not get convergence of \( L_\theta(s) \). On top of that, by the functional equation we see that \( \lim_{t \to 0} \frac{\theta(t)}{t^{1/2}} = 1 \), and so, for \( s \leq 0 \) the transform will also diverge since it will have order of magnitude \( t^{s-1/2} \) near \( t = 0 \). We overcome this problem by subtracting from the theta function, the function \( g \) which satisfies \( g(t) = t^{-1/2} \), for \( t < 1 \) and \( g = 1 \), for \( t \geq 1 \). We have the functional equation

\[
H_{1/2,1}(g) = g \quad \text{and so} \quad H_{1/2,1}(\theta - g) = \theta - g.
\]

From the definition of \( \theta \) we see that for \( t \geq 1 \) we have

\[
\theta - g = 2 \sum_{n=1}^{\infty} e^{-\pi n^2 t}
\]

which decays exponentially as \( t \to \infty \). By the functional equation the same is true as \( t \to 0 \).

We relate \( \zeta(s) \) to \( L_{\theta-g}(s/2) \). We will break up the Mellin transform into two parts

\[
L_{\theta-g}(s/2) = \int_0^1 t^{s/2} (\theta(t) - 1) \frac{dt}{t} + \int_1^\infty t^{s/2}(\theta(t) - 1) \frac{dt}{t}
\]

Assuming \( \Re(s) > 1 \) and considering separately the constant term of \( \theta \) in the first summand we can make the following evaluation

\[
L_{\theta-g}(s/2) = 2 \sum_{n \geq 1} \int_0^1 t^{s/2} e^{-\pi n^2 t} \frac{dt}{t} + \int_0^1 t^{s/2-1} dt - \int_0^1 t^{s/2-3/2} dt
\]

\[
+2 \int_1^\infty \sum_{n \geq 1} \int_0^1 t^{s/2} e^{-\pi n^2 t} \frac{dt}{t}
\]

\[
= 2(\sum_{n \geq 1} \int_0^\infty t^{s/2} e^{-\pi n^2 t} \frac{dt}{t} - 1/s - 1/s)
\]

By the change of variable \( u = \pi n^2 t \) we have

\[
\int_0^\infty t^{s/2} e^{-\pi n^2 t} \frac{dt}{t} = (\pi n^2)^{-s/2} \int_0^\infty u^{s/2} e^{-u} \frac{du}{u} = (\pi)^{-s/2} n^{-s} \Gamma(s/2)
\]

where by definition, the Gamma function is the Mellin transform of the exponential function \( e^{-t} \). The Gamma function is initially defined for \( \Re(s) > 0 \) because of convergence of the integral near \( t = 0 \). Using integration by parts it is easy to check that \( \Gamma(s+1) = s \Gamma(s) \) or \( \Gamma(s) = \Gamma(s+1)/s \). This allows us to extend the definition to all values of \( s \), with poles at \( 0, -1, -2, \ldots \). Using some more properties of the Gamma function it can be shown that \( 1/\Gamma(s) \)
is defined for all complex values of $s$ and that $\lim_{s \to 0} \frac{1}{s^{1/(s)}} = 1$. Summing equation (78) for all $n$ we conclude that

$$L_{\theta - g}(s/2) = 2(\pi^{-s/2}\Gamma(s/2)\zeta(s) - \frac{1}{s} - \frac{1}{1-s})$$

for $Re(s) > 1$. Since all the functions involved in the formula are defined (some with poles) in the whole plane, we can use it to define an analytic continuation for $\zeta(s)$ with a single pole at $s = 1$. Moreover, since $\theta - g$ has weight $1/2$ and $N = 1$ we have $L_{\theta - g}(s/2) = L_{\theta - g}(1/2 - s/2) = L_{\theta - g}((1 - s)/2)$. Since the sum $\frac{1}{s} + \frac{1}{s - 1}$ also is preserved by $s \to 1 - s$ we see that

(79) $\Lambda(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$

also satisfies

(80) $\Lambda(s) = \Lambda(1 - s)$

This is the functional equation for the Riemann zeta function. Having described Riemann’s use of the invariance of the theta function under the involution $H_{1/2,1}$, we turn to some basic queueing theory which is needed in the analysis of the mini-market queues.

2. Basic queueing theory

2.1. Little’s law. Consider a rather general queueing system where we make no assumptions on the arrival times and job-sizes of customers. We allow customers to enter the system at any time after time $t = 0$. If we let $Q(t)$ be the number of customers in the system at time $t$, then the average number of customers in the queue during a time interval $[0, t]$ is

$$E_t(Q) = \frac{\int_0^t Q(\tau)d\tau}{t}$$

Similarly, we may consider the function $A(t)$, the number of customers that arrived during time interval $[0, t]$.

The average arrival rate $\lambda_t$ of customers during the time interval $[0, t]$ is given by

$$\lambda_t = A(t)/t$$

The average waiting time $E_t(W)$ in the system during the time interval $[0, t]$ is the total time spent in the system by all customers, divided by the number of customers that entered the system. But, notice that the total amount of time spent by customers in the system is nothing but $\int_0^t Q(\tau)d\tau$, so we have

$$E_t(W) = \frac{\int_0^t Q(\tau)d\tau}{A(t)}.$$ From these definitions we have Little’s law, [122],

$$E_t(Q) = \lambda_t E_t(W)$$

In the systems that we are interested in the queues are described by stochastic processes. We will only consider cases where with probability 1, over all realizations (instances, “system runs”) the values of $E_t(Q)$, $\lambda_t$ and $E_t(W)$ tend as $t \to \infty$ to a constant limit which is independent of the particular
instance of the stochastic process. In such systems we define $E(W), E(Q)$ and $\lambda$ as the limits as $t \to \infty$ of $E_t(W), E_t(Q)$ and $\lambda_t$ respectively. Little's law states that if these limits exist then they must satisfy

\[ E(Q) = \lambda E(W) \]

2.2. A restricted class of queues and some notation. To make further progress we need to restrict ourselves to processes which are more manageable, like the ones that appear in the mini-market problem. We will assume that inter-arrival times between consecutive customers arriving at the queues, are independent of each other and identically distributed. We will also assume that the job sizes of customers are independent of each other and of arrival times and are again identically distributed.

In 1953, D. Kendall introduced some award winning notation to describe queueing systems, which has become standard, \cite{109}. In this notation, a general distribution is denoted by the letter "G", an exponential distribution by the letter "M", probably for the Markov property that it possesses.

When we describe a queue, we first specify the inter-arrival time distribution, followed by the job size distribution, followed by the number of servers, followed by the queue management discipline. For example, if the inter-arrival distribution is exponential, the job size distribution is general, there are $h$ servers and the customers are serviced in First come First served order (FIFO), then we denote the queue by $M/G/h/FIFO$, or, since FIFO is assumed by default, by $M/G/h$. If under the same assumptions customers re-ordered so that the customer with shortest job will be served next (SJF), we designate the queue by $M/G/h/SJF$.

The $M/G/h$ queue where inter-arrivals are memoryless can not be analyzed in closed form, and one needs to settle for some approximations.

Luckily the $M/G/1$ queue can be analyzed in a satisfactory manner via the Pollaczek-Khinchine formula, \cite{144, 111}, which we proceed to develop. Once we have that formula we can proceed to deal with express lines in the mini-market, which may be denoted by $M/G/2/Exp$. The analysis uses a few facts on Markov chains that can be found in the appendix.

2.3. The Pollaczek-khinchine formula. We will use the method of embedded Markov chains that was first developed in \cite{108}, To analyze an $M/G/1$ queue.

The idea is to consider the number of customers in the system at the time of departure of the $n$’th customer. The Markov chain that we will consider is $Q_i$, the number of customers left in the system following the departure of the $i$’th customer upon completion of service. Let $v_j$ denote the number of customers which arrive during the service of the $j$’th customer. If $Q_i > 0$ then the customers which will be left in the system following the departure of the customer $i + 1$ will be those which arrive while customer $i + 1$ is being served, together with those which were left behind at the departure of the $i$’th customer, excluding customer $i + 1$ which leaves the system. This leads
to the relation
\[(82) \quad Q_{i+1} = Q_i + v_{i+1} - 1\]
if \(Q_i > 0\). If no customers are left upon the departure of customer \(i\), namely, \(Q_i = 0\), then customer \(i + 1\) has to arrive first, and we will be left upon his departure with all the customers which arrived during his/her service time. This leads to the equation
\[(83) \quad Q_{i+1} = v_{i+1}\]
if \(Q_i = 0\). We consider the expression \(v_{i+1}\). Let \(Y(t)\) be the random variable describing the number of arrivals during an interval of length \(t\). Since we assume Poisson arrivals with arrival rate \(\lambda\) we have
\[(84) \quad P_r(Y(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}\]
We can now compute the probabilities \(\alpha_k = P_r(v_j = k)\). We first observe that they are independent of \(j\), the customer index. This is because the number of arrivals depends only on the service time (the memoryless property of the arrival process) and the service times were i.i.d.

Conditioning on the length of service we get from our discussion so far that
\[(85) \quad \alpha_k = \int_0^\infty \frac{(\lambda t)^k}{k!} e^{-\lambda t} f(t) dt\]
where \(f(t)\) is the density of the service time distribution.

Let us denote the Laplace transform, see the appendix, of the job size distribution by \(B_X\), and let
\[\alpha(z) = \sum_{k=0}^{\infty} \alpha_k z^k\]
be the generating function corresponding to the \(\alpha_k\). It follows from equation (85) that
\[(86) \quad \alpha(z) = B_X(\lambda - \lambda z)\]
Using the computed values of \(\alpha_k\) we see from equations (82,83) that the \(Q_i\) form indeed a Markov process. It is also clear that all the states of the process are aperiodic, since the probability of remaining in the same state is positive, since all \(\alpha_k\) are positive. It is also easy to see that the chain is irreducible. If \(Q_i = m\) then with positive probability \(Q_{i+1} = m + k\) for any \(k \geq 0\). Also \(Q_{i+1} = Q_i - 1\) with probability \(\alpha_0 > 0\), which means that after \(k\) time steps we reach the state \(m - k\) with probability \(\alpha_0^k > 0\).

We now explain under what condition we expect the process to be ergodic. We show that the state 0 is ergodic with finite average return time. Notice that excluding the state \(m = 0\), the probability of moving from state \(m\) to state \(l\) depends only on the difference \(m - l\). We may assign to our Markov process starting at state 0 a branching process as in chapter 4 with birth distribution \(Z = v\), which is given by the \(\alpha_k\). We notice that the size of
2. BASIC QUEUING THEORY

the population after each birth-death event is Markovian with a law that is identical with the law of our queue size Markov process as long as the queue is non empty. Returning to the state 0 means that the branching process became extinct. Under this equivalence, the number of time steps required to return to state 0 is simply the size (number of nodes) of the branching process, plus 1 (the initial step). The average size of the process is the sum of average sizes of the \( k \)’th generation, which according to (66) is given by \( E(v)^k \), hence the average return time to state 0 is \( \frac{1}{1-E(v)} \). We conclude that the state is ergodic if and only if \( E(v) < 1 \). We also note that this shows that \( q_0 = 1 - E(v) \). The quantity \( 1 - q_0 = E(v) \) represents the probability of finding customers in the system, or stated differently, the probability that the system is being utilized (is doing work). Consequently, this quantity is the utilization \( \rho = \lambda E(X) \). We note that the case \( E(v) = 1 \) provides an example of a persistent null chain.

Assuming that \( \rho < 1 \) we know that there is a limiting distribution for \( Q_i \) which we denote by \( Q_i \). It can be shown that in fact all moments of \( Q_i \) converge to the moments of \( Q \). It can also be shown that \( Q \) is the limiting distribution of queue length for \( t \to \infty \), not just the times when a job finishes. The reason for the latter is the memoryless property of the arrival process, see [113] for more details.

Let \( q_j \) be the probability of having \( j \) customers in the limiting distribution \( Q \). According to equations (208,82,82) the \( q_j \) and \( \alpha_k \) must satisfy the relation

\[
q_j = \sum_{k=0}^{j} q_{j+1-k} \alpha_k + q_0 \alpha_j = \sum_{k=0}^{j} q_{j+1-k} \alpha_k + (1 - \rho) \alpha_j
\]

In terms of the generating functions \( q(z) = \sum_{j=0}^{\infty} q_j z^j \) and \( \alpha(z) \) this is equivalent after some simple manipulations to the relation

\[
q(z) = \frac{(q(z) - q_0) \alpha(z)}{z} + (1 - \rho) \alpha(z) = \frac{(1 - \rho)(1 - z) \alpha(z)}{\alpha(z) - z}
\]

We recall that the average queue length is given by \( q'(1) \). Differentiating (87) we obtain

\[
q'(z) = (1 - \rho) \frac{\alpha'(z) z (z - 1) - \alpha(z)(\alpha(z) - 1)}{(\alpha(z) - z)^2}
\]

Applying L’Hospital at \( z = 1 \) we obtain

\[
q'(1) = (1 - \rho) \lim_{z \to 1} \frac{\alpha''(z) z (z - 1) + 2 \alpha'(z)(z - \alpha(z))}{2(\alpha(z) - z)(\alpha'(z) - 1)}
\]

\[
= (1 - \rho) \lim_{z \to 1} \frac{\alpha''(z) z (z - 1) + 2(\alpha(z) - z)(\alpha'(z) - 1)}{2(\alpha(z) - z)(\alpha'(z) - 1)} - \frac{\alpha'(z)}{\alpha'(z) - 1}
\]

By (86) and the appendix, we have

\[
\alpha'(1) = -\lambda B'(0) = \lambda E(X) = \rho
\]
Plugging this expression we get

\[ q'(1) = \lambda E(X) - \lim_{z \to 1} \frac{\alpha''(z)(z - 1)}{2(\alpha(z) - z)} \]

Applying L'Hospital again for the second summand we obtain

\[ q'(1) = \lambda E(X) - \lim_{z \to 1} \frac{\alpha'''(z)(z - 1) + \alpha''(z)(2z - 1)}{2(\alpha'(z) - 1)} = \lambda E(X) + \frac{\alpha''(1)}{2(1 - \rho)} \]

Finally, since \( \alpha''(1) = \lambda^2 B_X''(0) = \lambda^2 E(X^2) \), we obtain the formula

\[ E(Q) = \lambda E(X) + \frac{\lambda^2 E(X^2)}{2(1 - \rho)} \]

Since we have included in \( Q \) the customer which is currently getting service, by Little's law we have \( E(Q)/\lambda = E(W_{\text{tot}}) \) where \( W_{\text{tot}} \) refers to total time in the system. We will refer to the proper waiting time which does not include the service time by \( W \). We conclude that

\[ E(W_{\text{tot}}) = E(X) + \frac{\lambda E(X^2)}{2(1 - \rho)} \]

taking out the average service time we arrive at the equation

\[ E(W) = \frac{\lambda E(X^2)}{2(1 - \rho)} \]

This is the classic Pollaczek-Khinchine formula for the mean waiting time, which was proved independently by F. Pollaczek, [144], and A. Khinchine, [111].

Still, \( E(W) \) is not a perfect measure for performance. The reason is that it changes if we change the time units of measurement. If we measure time in seconds rather than minutes, the average waiting time will increase 60 fold. We would like an invariant performance measure. Therefore, we introduce the average normalized waiting time

\[ E(N) = E(W)/E(X) \]

Since a change in time units will affect proportionately both \( E(W) \) and \( E(X) \) we get an invariant performance measure. In our case we get

\[ E(N) = \frac{\rho E(X^2)}{2(1 - \rho) E(X)^2} \]

Notice that we have separated \( E(N) \) multiplicatively into two separate components, a component that depends only on the utilization

\[ E(N) = \frac{\rho}{2(1 - \rho)} \]
and a component that depends only on the service time distribution

\[ E(N)_X = \frac{E(X^2)}{E(X)^2} \tag{93} \]

the last component also being known as the \textbf{squared coefficient of variation} and denoted by \( C_X^2 \).

\section*{2.4. Involutions, projective densities and target functions.}

We would like to consider the involutions \( H_{k,N} \) as acting on density functions. Density functions satisfy two conditions. They take non-negative real values. This property is certainly preserved by \( H_{k,N} \). We also have the total probability normalization property

\[ \int_0^\infty f(t)dt = 1 \]

This property is not preserved by \( H_{k,N} \). To overcome this difficulty we introduce the following definition.

**Definition 99.** A non-negative function \( f \), for which \( 0 < \int_0^\infty t^s f(t)\frac{dt}{t} < \infty \) for \( 1 \leq s \leq k - 1 \) is said to be a \( k \)-finite density.

By lemma 16, \( k \)-finite densities are invariant under \( H_{k,N} \). If \( f \) is such a function we can multiply it by the constant \( c_f = 1/(\int_0^\infty f(t)dt) \) to obtain a density. In fact \( c_ff \) is the only multiple of \( f \) which is a density. Thus we are interested in non-negative non zero functions up to a constant multiple. It would then be useful if we can state queueing theoretic quantities of interest such as average waiting time in formulas which are invariant with respect to multiplication of the density function by a constant.

As a first step we would like to replace expressions such as \( E(X^m) \) which assume a normalized density. By definition, if \( f \) denotes the density associated with a random variable \( X \) we have

\[ E(X^m) = L_f(m + 1) \]

for all \( m \). We can therefore write \( E(N)_X \) in terms of the Mellin transform as

\[ E(N)_X = \frac{L_f(3)}{L_f(2)^2} \tag{94} \]

However, if we replace \( f \) by a constant multiple \( cf \) on the left hand side, the value will be multiplied by \( 1/c \). We would like an expression which does not change if we replace \( f \) by \( cf \) and which on densities coincides with the expression in equation (94). The procedure for doing so is rather trivial, and is inspired by the way functions on projective space are constructed in algebraic geometry.

We note that for a distribution \( X \) we have \( E(X^0) = L_f(1) = 1 \), Plugging this into our Mellin transform expression for \( E_N(X) \) we have

\[ E(N)_X = \frac{L_f(1)L_f(3)}{L_f(2)^2} \tag{95} \]
This expression has the advantage that it remains the same if we plug into it a density \( f \) or a constant multiple \( cf \). The requirement that the formula will be invariant under multiplication by a constant translates into the condition that the number of Mellin transforms in the numerator and denominator should be equal. The condition that the formula should be invariant under change of time units translates into the condition that the sum of values at which the Mellin transform is evaluated is equal in the numerator and the denominator.

We can apply a similar procedure to other popular target functions.

**Definition 100.** Given a job which waited time \( w \) and whose size is \( x \) we define the **slowdown** as \( w/x \).

Average slowdown is a popular measure of a system’s performance. In a FIFO queue, there is no relation between the waiting time of a job and its size, so \( E(S) = E(W)E(X^{-1}) \), where \( S \) denotes the slowdown. This leads to the formula

\[
E(S) = \frac{\rho}{2(1-\rho)} \frac{L_f(0)L_f(3)}{L_f(2)}
\]

which after our homogenization procedure leads to

\[
E(S) = \frac{\rho}{2(1-\rho)} \frac{L_f(0)L_f(3)}{L_f(1)L_f(2)}
\]

There are also formulas for higher moments of the waiting time \( W \). As before \( E(W^k) \) is not a time unit invariant quantity. To normalize we consider the random variable \( N_m = \frac{W^m}{E(X^m)} \). For \( E(N_m) \), the normalized \( m \)’th moment of waiting time, we have a recursive formula due to Takacs, [168]

\[
E(N_m) = \frac{\rho}{1-\rho} \frac{1}{E(X)E(X^m)} \sum_{i=1}^{m} \frac{B_{m,i}}{i+1} E(X^{i+1})E(X^{m-i})E(N_{m-i})
\]

where \( B_{m,i} \) is the binomial coefficient and \( E(N_0) \) is 1 by convention. We can see by induction that the expressions for \( E(N_m) \) will be invariant both for multiplication of the density by a constant and for a change of time units. We also see from this formula that all moments of waiting time are combinations of the Mellin transform evaluated at integer points.

### 3. Variants of the M/G/1 queue

Apart from the basic case of an M/G/1 queue, we can analyze closely related queues by applying variants of the P-K formula. In this section we explore a couple of such queues including the mini-market scenario.
3.1. SITA policies. One important consequence of the P-K formula is that large variance (second moment) in the job size distribution leads to large waiting times. This motivates the notion of "express lines" with cutoffs as we have encountered in the mini-market problem. In such queueing systems each host is responsible only for jobs in a certain range. This reduces the variance at the host. We now introduce SITA policies, which formalize the notion of express lines. These policies have been introduced in the computer science setting in [86].

**Definition 101.** A SITA policy can be described as follows:
Consider a system with $h$ coupled hosts, numbered $1, \ldots, h$. Let $s_1 < s_2 < \ldots < s_{h-1}$ be a set of cutoff parameters which are fixed. We let $s_0 = 0$ and $s_h = \infty$. Let $\sigma$ be a permutation on the set of hosts $1, \ldots, h$. We assume that job-sizes are known. An incoming job of size $s$ is dispatched to host $\sigma(i)$, such that $s_{i-1} \leq s < s_i$. Requests are serviced in each host in a first come, first served (FIFO) order.

Assume that the arrivals to the SITA system are Poisson. We assume for now that the hosts are identical and we let $X$ be the generic job-size distribution. Let us extend the P-K formula to cover SITA systems with a given set of cutoffs $s_1 < s_2 < \ldots < s_{h-1}$. Let $I = (a, b]$ be some interval with $a, b \geq 0$. We let $X_I$ be the distribution of job-sizes in the interval $I$. If $X$ is given by some density $f$, then the restriction of $f$ to the interval $I$ is, up to a constant multiple, the density for $X_I$. In terms of the incomplete Mellin transform $L_{f,I}(s) = \int_I f(t) t^{s-1} dt$ we have

$$E(X_I^k) = \frac{L_{f,I}(k + 1)}{L_{f,I}(1)}$$

Let $I_j = (s_{j-1}, s_j]$. Let $E(W_j)$ be the average waiting time (service not included) of jobs in the $j$'th queue of a SITA system. If $f$ is a proper density, namely, $L_f(1) = 1$, then the portion of jobs in the $j$'th queue is $L_{f,I_j}(1)$.

We conclude that $E(W) = \sum_j E(W_j)$.

First we compute the rate of arrival of jobs at the $j$'th host

$$\lambda_j = \lambda L_{f,I_j}(1) = \rho \frac{L_{f,I_j}(1)}{L_f(2)}$$

From this we can compute the utilization at the $j$'th server

$$\rho_j = \lambda_j E(X_j) = \rho \frac{L_{f,I_j}(1) L_{f,I_j}(2)}{L_f(2) L_{f,I_j}(1)} = \rho \frac{L_{f,I_j}(2)}{L_f(2)}$$

Combining with (91) and (99) we see that the average waiting time at host $j$ is given by

$$E(W_j) = \rho \frac{L_{f,I_j}(1) L_{f,I_j}(3)}{L_f(2) L_{f,I_j}(1)} (2(1 - \rho \frac{L_{f,I_j}(2)}{L_f(2)}))^{-1}$$
\[ E_j(N) = \frac{\rho}{2} (1 - \rho \frac{L_{f,I_j}(2)}{L_f(2)})^{-1} \frac{L_{f,I_j}(1)}{L_f(2)} \]

To get the weighted contribution of host \( j \), to the average normalized waiting time \( E(N) \), we need to multiply by \( L_{f,I_j}(1) \), the portion of jobs arriving at host \( j \), and to divide by \( L_f(2) = E(X) \). Calling the weighted contribution of host \( j \) to \( E(N) \) by \( E_j(N) \) we have

\[ E_j(N) = \frac{\rho}{2} (1 - \rho \frac{L_{f,I_j}(2)}{L_f(2)})^{-1} \frac{L_{f,I_j}(1) L_{f,I_j}(3)}{L_f(2)^2} \]

and

\[ E(N) = \sum_j E_j(N) \]

Starting with (96) we can also develop the analogous formula for the contribution of the \( j \)'th host to average slowdown in a SITA queue

\[ E_j(S) = \frac{\rho}{2} (1 - \rho \frac{L_{f,I_j}(2)}{L_f(2)})^{-1} \frac{L_{f,I_j}(0) L_{f,I_j}(3)}{L_{f,I_j}(1) L_{f,I_j}(2)} \]

### 3.2. Priority queues.

**Definition 102.** A priority queue is a queueing system where the customer to be serviced next, is chosen based upon their arrival time, job size, group membership or some combination of the above. Examples with suggestive names include, First In First Out (FIFO), Last In First Out (LIFO), Shortest Job First (SJF), Longest Job First (LJF) and many others.

We will assume that the system has a single server and Poisson arrivals. Our discussion closely follows [113] volume II.

We will assume that each arriving customer belongs to a priority class which will be denoted by \( p \). The set of all priority classes is assumed to be a finite set \( \{1, \ldots, P\} \). The generalization to a sequence or continuum of classes will be simple to obtain via a limiting approximation process.

We will assume that a higher priority index leads to preferential treatment in the queue. In some cases the priority may be dynamical, i.e., may change with time, however, we will restrict our attention to time independent priorities. We will also restrict the discussion to the non-preemptive case which means that service to a customer is not interrupted.

We assume that customers of class \( p \) have Poisson arrival with rate \( \lambda_p \) and that their job size is sampled i.i.d. from a distribution \( X_p \). We have the following basic equations for the total arrival rate \( \lambda \), average job-size moments \( E(X^j) \), class \( p \) load \( \rho_p \), and total load \( \rho \)

\[ \lambda = \sum_p \lambda_p \]

\[ E(X^j) = \sum_p \frac{\lambda_p}{\lambda} E(X_p^j) \]
\[ \rho_p = \lambda_p E(X_p) \]

\[ \rho = \sum_p \rho_p \]

Let \( W_p \) denote the average waiting time (service not included) of a job with priority \( p \). Let \( W_0 \) be the amount of time that an arriving customer has to wait until the job currently being processed finishes. This waiting time is independent of the class \( p \), since we are assuming independent Poisson arrivals.

Consider a customer that arrives at some random point in time. With probability \( 1 - \rho \) it will arrive at an empty queue. let \( \bar{W}_0 \) be the average waiting time until the current job finishes assuming that there is a job being processed. Since the arrival process is Poisson (no memory or dependence on prior arrivals), on average, the arrival will be at the middle of the service period of the present customer. This is part of a general principal that in queueing theory terminology is sometimes referred to as PASTA (Poisson Arrivals See Time Averages).

It is more likely that the arrival occurred during a long service period rather than a short service period. The portion of jobs that have size between \( t \) and \( t + dt \) is approximately \( f(t) dt \). Since all such jobs take approximately \( t \) processing time, the portion of processing time that such jobs take is proportional to \( f(t)t \) and if such a job is being processed, the arrival will wait on average \( t/2 \). To normalize \( tf(t) \) to a density we must divide by \( E(X) = \int_0^\infty tf(t) dt \).

**Definition 103.** The **residual job-size density** is the normalized density \( tf(t)/E(X) \) which measures the portion of processing time of jobs of size \( t \) out of all jobs.

A customer which arrives while another customer is being processed will wait on average \( \int_0^\infty \frac{1}{2} tf(t) dt / E(X) = E(X^2)/2E(X) \) time for that customer to complete service. Since we find a job being processed with probability \( \rho \) and otherwise the arrival does not wait at all we find that

\[ W_0 = \rho \frac{E(X^2)}{2E(X)} = \rho \frac{L_f(3)}{2L_f(2)} \]

This looks a lot like the P-K formula and indeed, we could use these arguments to derive the P-K formula for average normalized waiting time. In terms of the classes we have

\[ W_0 = \sum_p \rho_p \frac{E(X_p^2)}{2E(X_p)} = \frac{1}{2} \sum_p \lambda_p E(X_p^2) \]

We note that by (108), \( W_0 \) depends only on the total job-size distribution.
Assume that a customer belongs to group $p$. Let $Q_i$ be the number of customers of group $i$ which are in the queue at the time of arrival of the customer. The customers with priority index $i \geq p$ will receive service before our customer. This will lead to an average delay of
\begin{equation}
\sum_{i \geq p} E(Q_i) E(X_i)
\end{equation}

In addition, let $M_i$ be the number of customers of priority $i$ which arrive while our customer is waiting in line to get service. If a customer has priority $i > p$ and arrives while our customer is still in line it will be serviced before our customer adding an average delay of
\begin{equation}
\sum_{i > p} E(M_i) E(X_i)
\end{equation}

By Little’s law we have
\begin{equation}
E(Q_i) = \lambda W_i
\end{equation}

Since the arrivals of each priority class are independent Poisson and the average waiting time of a customer of class $p$ is by definition $W_p$ we get
\begin{equation}
E(M_i) = \lambda_i W_p
\end{equation}

Putting it all together we get a system of equations
\begin{equation}
W_p = W_0 + \sum_{i \geq p} E(X_i) \lambda_i W_i + \sum_{i > p} E(X_i) \lambda_i W_p
\end{equation}

Moving the terms involving $W_p$ to the left and using equation (106) we get
\begin{equation}
W_p = \frac{W_0 + \sum_{i > p} \rho_i W_i}{1 - \sum_{i \geq p} \rho_i}
\end{equation}

This expresses $W_p$ in terms of $W_i$, $i > p$ so we can solve recursively starting with $W_P$. Letting $\sigma_p = \sum_{i \geq p} \rho_i$ the solution to the recursion is
\begin{equation}
W_p = \frac{W_0}{(1 - \sigma_p)(1 - \sigma_{p+1})}
\end{equation}

We note that applying this to a system where there is only a single class $p = 1$ we get the P-K formula for mean waiting time.

We can apply the result to obtain the average waiting time in the Shortest Job First (SJF) policy. In SJF, each job size $t$ is a class and unlike our convention so far it has priority over all jobs of class $t' > t$. The distribution $X_t$ is the deterministic distribution which assigns probability 1 to the value $t$. We can analyze this system by discretizing it and taking the limit. Discretization is obtained by grouping together all jobs whose size is in the range $[jdt, (j+1)dt]$ for $j \leq P$. Here $dt$ is small and $P$ is very large. We also group all jobs of size $t > Pdt$ into a single class. Letting $t = jdt$ the arrival
rate $\lambda_j$ will be approximately $f(t)dt$. In the limit, sums will be replaced by the corresponding integrals and following the definitions we get

\begin{equation}
W_t = \frac{W_0}{(1 - \rho_t)^2}
\end{equation}

where

\begin{equation}
\rho_t = \int_0^t \tau f(\tau) d\tau = L_{f,t}(2)
\end{equation}

with $I_t = (0, t)$.

Similarly, for the Longest Job First (LJF) policy we have

\begin{equation}
W_t = \frac{W_0}{(1 - \rho_t')^2}
\end{equation}

where

\begin{equation}
\rho_t' = \int_t^\infty \tau f(\tau) d\tau = L_{f,t'}(2)
\end{equation}

with $I'_t = (t, \infty)$.

Let $N_{f,SJF}(t) = W_t/E(X)$ be the average normalized waiting time of a job of size $t$, when the policy is SJF and the job size density is $f$ and similarly $N_{f,LJF}(t)$ the corresponding quantity in a LJF managed queue. Putting together equations (108,117,120) we can express $N_{f,SJF}$ in terms of incomplete Mellin transforms as

\begin{equation}
N_{f,SJF}(t) = \frac{\rho}{2(1 - \rho L_{f,t}(2)^2)} \frac{L_f(1)L_f(3)}{(L_f(2))^2}
\end{equation}

and similarly for LJF we have

\begin{equation}
N_{f,LJF}(t) = \frac{\rho}{2(1 - \rho L_{f,t'}(2)^2)} \frac{L_f(1)L_f(3)}{(L_f(2))^2}
\end{equation}

4. Duality theory

We are in a position to bring together the analysis via the P-K formula of SITA, SJF and LJF queues with Riemann’s lemma on the behavior of the Mellin transform w.r.t. the involutions $H_{k,n}$. The basic result is the following, see [73] and [20].

**Theorem 55.** Let $f$ be a 4-finite density. For an M/G/1 queue with fixed utilization $\rho$ and job size density $f$, we have, for any level $n > 0$,

\begin{equation}
E(N)(f, \rho) = E(N)(H_{4,n}(f))
\end{equation}

In addition,

\begin{equation}
H_{0,n}(N_{f,SJF}) = N_{H_{4,n}(f),LJF}
\end{equation}

For densities such that $f(t)/t$ is 3-finite we have

\begin{equation}
E(S)(f) = E(S)(H_{3,n}(f))
\end{equation}
Consider a SITA queue with \( h \) linearly coupled hosts with speeds given by the vector \( b = (b_1, \ldots, b_h) \), cutoffs \( s_1, \ldots, s_{h-1} \), utilization \( \rho \), job size density function \( f \) and a permutation \( \sigma \), where host \( \sigma(i) \) is assigned to handle jobs whose size \( s \) is in the interval \([s_{i-1}, s_i)\).

Let \( E(N)(f, h, \rho, b, \sigma, s_1, \ldots, s_{h-1}) \) denote the average normalized waiting time of the SITA queue.

Let \( \hat{s} = \frac{1}{\hat{m}_s} \). Then, the following dualities hold for any level \( n > 0 \),

\[
E(N)(f, h, \rho, b, \sigma, s_1, \ldots, s_{h-1}) = E(N)(H_{4,n}(f), h, \rho, b, \sigma, \hat{s}_1, \ldots, \hat{s}_1)
\]

**Proof:** The single host results follow from Riemann’s lemma and the P-K average normalized waiting time expressions for the various queues. For SITA queues, let \( I_j = (s_{j-1}, s_j) \). Then, in the notation of lemma 16, \( \hat{I}_j = (\hat{s}_{h-j}, \hat{s}_{h-j+1}) \). The host which is assigned to handle the \( j \)-th interval of the original queue, handles the \( h-j-1 \) interval in the dual system, i.e., it handles the dual interval. Denoting the weighted contribution to normalized average waiting time of the \( j \)-th host to the original queue by \( E_j(N) \) and to the dual queue by \( \hat{E}_j(\hat{N}) \), we have by equation (101) and lemma 16 that \( E_j(N) = \hat{E}_{h-j+1}(\hat{N}) \), which proves the assertion for SITA systems. *q.e.d.*

We remark that the first statement in the theorem is incorrect if we replace the normalized average waiting time \( E(N) \) by the average waiting time \( E(W) \) despite the fact that the expression is invariant under \( H_{4,n} \). The reason is that the expression for \( E(W) \) depends on working with densities which are normalized so that \( L_f(1) = 1 \), while \( H_{4,n} \) does not map such densities to themselves.

We consider the mini-market problem using duality. We need a result of Harchol-Balter and Vesilo, [88], that the optimal cutoff in a SITA system with two identical hosts is essentially unique.

**Theorem 56.** Let \( f \) be a density function for which \( L_{f,I_1}(2) \) is a strictly increasing function of \( t \). A SITA queue with job size density \( f > 0 \), load \( \rho \), and two identical hosts has a unique cutoff, \( s \), which minimizes average normalized waiting time.

**Proof:** Recall that \( I_i = (0, t] \). Let \( m_i(t) = L_{f,I_i}(i+1) \) be the \( i \)-th incomplete moment of \( f \). Using the assumption that \( f > 0 \) we can re-parametrize the time coordinate via \( m_1(t) \). From the definition we see that \( dm_i(t)/dt = tf(t) \) and so \( dm_i/m_1 = (dm_i/dt)/(dm_1/dt) = t^{-1} \). We consider the contribution of the first host to average waiting time. In terms of the \( m_i \), it is given by

\[
E_1(N) = \frac{1}{2(1-\rho)} \frac{m_0 m_2}{m_1^2} = \int_0^{\infty} tf(t) dt.
\]

We will show that as a function of \( m_1 \) this is a convex function. We need to show that \( \frac{d^2}{dm_1^2} E_1(N) \geq 0 \). After some tedious manipulations it boils down to showing that

\[
\frac{d^2}{dm_1^2} m_0 m_2 \geq 0.
\]

We have \( \frac{d}{dm_1} m_0 m_2 = \frac{m_2}{t} + tm_0 \).
Differentiating again and using $dt/dm_1 = 1/(dm_1/dt) = \frac{1}{tf(t)}$ we get
$$\frac{d^2}{d^2m_1}m_0m_2 = 2 + \left(m_0 - \frac{m_2}{t^2}\right)\frac{1}{tf(t)}$$
but, $m_0(t)t^2 = \int_0^t t^2 f(\tau)d\tau > 0 \int_0^t \tau^2 f(\tau)d\tau = m_2(t)$ which implies that the second summand is non-negative and therefore that the contribution of the first host is convex. For the contribution of the second host $E_2(N)$ we can use duality to show that it equals the contribution of the first host in the queue with the dual job size density $H_{4,s}(f)$. The parameter $m_1$ is replaced in this case by $m_1^f - m_1$. However, this does not change the second derivative so we conclude that this contribution is also convex and therefore $E(N)$ is convex as a function of $m_1$. Since a convex function with a strictly positive second derivative has a unique minimum over an interval we obtain the desired result. *q.e.d.*

**Definition 104.** We say that a density $f$ is **self-dual of weight** $k$ if for some level $n > 0$ we have $f = H_{k,n}(f)$. We say in such a case that $n$ is a **level of $f$ for weight $k$**.

If $f$ is normalized, $k \neq 2$ and $E(X^{k-2})$ exists, then applying lemma 16 with $s = 1$ we obtain $n = E(X^{k-2})^{\frac{k-2}{k}}$ so the level is determined by $f$ and $k$. We call it the **conductor of $f$** and denote it by $n_f = n_{f,k}$.

The following result shows how to solve the mini-market management problem for self-dual job-size densities of weight 4.

**Theorem 57.** If $f$, the job-size density of a mini-market queue is self-dual of weight $k = 4$ and $E(X^2) < \infty$, then:

1) The optimal cutoff $s_{opt}$ for minimizing average waiting time when $b = 1$ is
$$s_{opt} = 1/\sqrt{n_f} = \sqrt{E(X^2)}$$
This is also the queue length balanced solution.

2) The assignment of servers to counters is not important
$$E(W)_{opt}(f, \rho, b, \bar{\sigma}) = E(W)_{opt}(f, \rho, b, \sigma)$$

3) We have
$$s_{opt}(f, \rho, b, \sigma) = \frac{1}{n_fs_{opt}(f, \rho, b, \bar{\sigma})}$$

**Proof:** From theorem 55 it is clear that if $f$ is self-dual of weight 4 and the hosts are equally powerful, then, if $s_{opt}$ is an optimal cutoff, so is $1/(n_fs_{opt})$. By uniqueness we have that the two values are equal and we obtain the first part of the theorem. The second part and third part also follow directly from theorem 55. *q.e.d.*
5. Distributions in queueing theory

From theorem 57 we see that all that is left to show, in order to prove theorem 53, is that the functions \( f_{A,B} \) are self-dual of weight 4. While this is an extremely difficult result which is well beyond the level of this book, we will at least try to place it in a larger context.

Before we consider the functions \( f_{A,B} \) we consider some distributions which often appear in the queueing theory literature when modeling job-size distributions. We concentrate on distributions which have nice properties w.r.t. the involutions \( H_{k,N} \). Other classical distributions which do not have nice properties w.r.t. \( H_{k,N} \) are described in the notes to this chapter.

In some cases we equate density functions \( f \) with non-negative functions on the positive reals. In such cases it should be understood that actual densities are obtained from \( f \) by restricting to any interval \( I \) for which \( 0 < \int_I f(t) dt < \infty \), with a similar construction for \( k \)-finite densities. In addition, following our discussion so far, we will not distinguish between the functions \( f \) and \( cf \) when \( c > 0 \).

5.1. Pareto distributions.

Definition 105. A Pareto distribution (or Pareto law) with parameter \( \alpha \) has a density function \( f \) of the form \( f(t) = t^{-\alpha - 1} \). If the density is restricted to a finite interval \( I = [a, b] \), \( 0 < a < b < \infty \), we say that the distribution is Bounded Pareto and denote it by \( P_{\alpha,I} \).

Pareto distributions also have a discrete counterpart which has non-negative values which is known as Zipf’s distribution (or Zipf’s law).

Some reasons why Pareto and Zipf laws should be commonly observed is given in Terence Tao’s blog entry, [169], where it is explained that these distributions have some universality properties as limits, much in the same way that the normal distribution is universal for limits of sums of i.i.d. random variables.

Pareto distributions have a very important property, they are scale invariant. Assume \( c > 0 \). For an interval \( I = [a, b] \) denote by \( cI \) the scaled interval \([ca, cb] \), which may be obtained from \( I \) by scaling the time unit. As generalized densities \( f_{\alpha,I}(ct) = f_{\alpha,cI}(t) \), which means that the distributions \( P_{\alpha,I} \) and \( P_{\alpha,cI} \) are essentially the same, or that the Pareto distribution looks the same at all scales.

Given an interval \( I \) we can define its multiplicative length

\[
l(I) = \exp(\int_I \frac{dt}{t}) = b/a
\]

This measure is invariant under scaling and we let \( P_{\alpha,l} \) be the Pareto distribution with parameter \( \alpha \), restricted to the interval \([1, l] \) of multiplicative length \( l \). In most of the discussion, the specific interval will not be important, as the results will depend only on the multiplicative length \( l(I) \).

The following lemma says that Bounded Pareto distributions are self-dual as a family of distributions.
Lemma 17. The following equality holds

\[ H_{k,1/l}(f_{\alpha,l}) = f_{k-\alpha-2,l} \]

In particular \( f_{\alpha,l} \) is self-dual of weight \((k/2) - 1\).

**Proof:** First we note that on the level of intervals \( H_{k,1/l}([1,l]) = [1,l] \) so the interval of restriction is self-dual. We have

\[
H_{k,1/l}(f_{\alpha,l})(t) = N^{-k/2}t^{-\alpha-1}
\]

\[
= N^{-k/2}t^{-\alpha-1}l^{-(k-\alpha-2)-1} = cf_{k-\alpha-2,l}
\]

As required. The second assertion follows trivially. *q.e.d.*

For \( k = 4 \), which is associated by theorem 55 with normalized average waiting time, we get that \( f_{1,l} \) is self-dual and more generally \( f_{\alpha,l} \) and \( f_{2-\alpha,l} \) are dual. Similarly, for \( k = 3 \), which is associated with average slowdown we get that \( f_{1/2,l} \) is self-dual and more generally \( f_{\alpha,l} \) and \( f_{1-\alpha,l} \) are dual.

### 5.2. Double Pareto distributions

It has been observed that in many cases the distribution of short jobs and of long jobs differ, \([25, 135]\). Such phenomenon can be modeled using the following type of densities

**Definition 106.** A **double Pareto distribution** is a distribution with density function of the form

\[ f(s) = s^{-\alpha_1-1} \]

for \( s \geq c \) and

\[ f(s) = s^{-\alpha_2-1} \]

for \( s \leq c \), where \( c > 0 \) is some constant.

Taking \( N = 1/c^2 \) and \( k = \alpha_1 + \alpha_2 + 2 \) we see that the double Pareto is self dual w.r.t. \( H_{k,N} \). For example, take \( \alpha_1 = -1 \), which means that for \( s \geq c \), \( f(s) = 1 \) and let \( \alpha_2 = -1/2 \), which means that \( f(s) = s^{-1/2} \) for \( s \leq c \), then \( k = 1/2 \). In fact, we have subtracted this double Pareto from the theta function, in order to analytically extend the definition of the zeta function.

### 5.3. Log-normal distributions

The problem with double Pareto distributions is that they have a break point. A smoother choice is the following.

**Definition 107.** A **log-normal distribution** is a distribution with density

\[ f(t) = \frac{1}{\sigma\sqrt{2\pi}} t \exp\left(\frac{-(\log(t) - \mu)^2}{2\sigma^2}\right) \]

where \( \sigma, \mu > 0 \) are parameters.
As the name suggests, the distribution is obtained by plugging \( \log(t) \) as the variable in the formula for the normal distribution with average \( \mu \) and standard deviation \( \sigma \). The \( \frac{1}{\tau} \) term comes from \( d \log(t) = \frac{dt}{\tau} \).

In many cases a log-normal distribution is used for modeling the body (rather than tails) of a distribution, \([133, 134, 148, 135]\). These distributions can smoothly interpolate between the two tails of the double Pareto. From the point of view of duality The log-normal distribution \( f \) has a remarkable property, for any \( k \) there exists an \( N(k) \) such that \( f \) is self dual w.r.t. \( H_{k,N(k)} \). This has been observed in \([73]\). We consider more generally generalized densities of the form

\[
(130) \quad f(t) = \text{Exp}[a \log^2(t) + b \log(t)]
\]

which up to a constant multiple contain all products of a Pareto distribution and a log-normal distribution. After a short calculation we get that

\[
f(t) = ct^{-2a \log(N) + 2b} f(1/Nt)
\]

for some constant \( c \), hence the function has weight \( 2a \log(N) - 2b \), when \( N \) is the level. By playing with the value of \( \log(N) \) we see that we can get self-duality w.r.t. any weight as long as \( a \neq 0 \). We also see that for each \( N > 0 \) there is a \( k = k(N) \) such that density is self-dual w.r.t. \( H_{k(N),N} \), this holds even when \( a = 0 \). We claim that this property essentially characterizes these distributions.

**Theorem 58.** Let \( g \) be a measurable function such that for any \( N > 0 \) there exists \( k = k(N) \) such that \( g = H_{k(N),N}(g) \), then \( g \) is a constant multiple of a generalized log-normal function.

**Proof:** It is somewhat easier to follow the argument if we make the change of variable \( x = \log(t) \) and consider \( g \) as a function of \( x \). The mapping \( t \to \frac{1}{Nt} \) becomes \( x \to -\log(N) - x \). Recall that the reflection of the real numbers centered at a point \( d \) is given by \( x \to 2d - x \), hence denoting this map by \( I_d(x) \) and setting \( d = -\frac{1}{2} \log(N) \) we see that the relation \( g(t) = N^{-k/2} t^{-k} g(\frac{1}{Nt}) \) becomes

\[
(131) \quad g(x) = e^{dk} e^{-k} g(I_d(x))
\]

Consider \( N_1 = 1 \), \( N_2 = 1/e \), the corresponding values of \( d \) are \( d_1 = 0 \), and \( d_2 = 1/2 \). Assume that the weights of \( g \) w.r.t. these two values of \( N \) are \( k_1, k_2 \) respectively. We can choose a generalized log-normal density which will have these weights, namely, with \( b = -k_1/2 \) and \( a = -1/2(k_2 - k_1) \). Taking the ratio of \( g \) with this generalized log-normal we obtain a function \( h \) which has weight \( k_1 = k_2 = 0 \) w.r.t. 1 and 1/e respectively, or after the change of variable \( h(x) = h(I_0(x)) = h(I_{1/2}(x)) \). If we have two reflections \( I_{d_1} \) and \( I_{d_2} \) then their composition is \( x \to 2d_2 - (2d_1 - x) = 2(d_2 - d_1) + x \), hence we conclude that \( h \) has period 1. Consider the reflection \( I_{1/4} \) which corresponds to \( N = e^{-1/2} \). By our assumption on \( g \) and the properties of the log-normal function we know that \( h \) will satisfy (131) for \( d = 1/4 \) with
some value of $k$, say $k(1/4)$. Taking $x = -3/4$ we see that $h(-3/4) = e^{k(1/4)}h(5/4) = h(-3/4)$ from which we conclude that $k(1/4) = 0$ and hence that $h$ has period $1/2$ by taking $d_1 = 0$ and $d_2 = 1/4$. Repeating with $d = 1/8$ and so on we see that $h$ has period $2^{-l}$ for all integers $l \geq 0$. However, by theorem 74 of the appendix a measurable function which has arbitrarily small periods is a constant and consequently, $g$ is a generalized log-normal distribution. q.e.d.

Having considered a few examples which are commonly considered in queueing theory, we turn to examples from number theory, explaining how they arise.

6. Modular forms

Modular forms are very special functions with many functional equations. There are many excellent introductions to the subject, please consult [160, 45, 119, 62, 114], among others. A less technical and proof-less sketch of closely related material can be found in [50], which uses a different application (to distributing points on the sphere) as an excuse to discuss this wonderful piece of mathematics.

Due to shortage of space and the wealth of these excellent sources our presentation is sketchy at best. We merely give hints for the arguments involved.

6.1. Modular forms for $SL_2(\mathbb{Z})$. The group $GL_2(\mathbb{R})^+$ of real $2 \times 2$ matrices with positive determinant acts on the complex upper half plane, $H$, consisting of complex numbers $z = x + yi$ such that $y > 0$. The action of $\gamma \in GL_2(\mathbb{R})^+$ is given by

\begin{equation}
\gamma(z) = (az + b)/(cz + d)
\end{equation}

where the matrix $\gamma$ is given by

\begin{equation}
\begin{pmatrix}
a & b \\
c & d
\end{pmatrix}
\end{equation}

It is easy to check that for such matrices $\gamma$ we have

\begin{equation}
Im(\gamma(z)) = det(\gamma) \frac{Im(z)}{|cz + d|^2}
\end{equation}

which also verifies that $H$ is indeed preserved by this mapping. In addition it is easy to check that this is a group action, namely

\begin{equation}
\gamma_1(\gamma_2(z)) = (\gamma_1\gamma_2)(z)
\end{equation}

**Definition 108.** We define the weight $k$ action of $\gamma \in GL_2(\mathbb{R})^+$ on a function $f$ defined on $H$ by the formula

\begin{equation}
f_{\gamma,k}(z) = det(\gamma)^k(az + b)^{-k}f(\gamma(z))
\end{equation}
It is easy to verify using (135) that this is also a group action

\[ f_{\gamma_1 \gamma_2, k} = (f_{\gamma_2, k})_{\gamma_1, k} \]

We denote by \( SL_2(\mathbb{Z}) \), the group of matrices with integer coefficients and determinant 1. For reasons that will become clear later on we will also denote \( SL_2(\mathbb{Z}) \) by \( \Gamma(1) \), or if no confusion arises by \( \Gamma \).

We will be interested in functions on \( \mathbb{H} \) which behave nicely w.r.t. the action of \( SL_2(\mathbb{Z}) \), but before discussing the functions we consider a few more basic facts about the action on \( \mathbb{H} \). The group \( SL_2(\mathbb{Z}) \) is generated by two elements, the matrix \( T \) given by

\[
\begin{pmatrix}
1 & 1 \\
0 & 1
\end{pmatrix}
\]

whose action on \( \mathbb{H} \) is given by

\[ T(z) = z + 1 \]

and the matrix \( S \), given by

\[
\begin{pmatrix}
0 & 1 \\
-1 & 0
\end{pmatrix}
\]

whose action on \( \mathbb{H} \) is given by

\[ S(z) = \frac{-1}{z} \]

Definition 109. A **fundamental domain** for the action of \( \gamma \) on \( \mathbb{H} \) is a (simply) connected set \( F \subset \mathbb{H} \) such that for any \( z \in \mathbb{H} \) there exists a \( \gamma \in \Gamma \) such that \( \gamma(z) \in F \) and that if \( z_1, z_2 \) are in the interior of \( F \) then they are not mapped to each other by \( \Gamma \).

Consider the domain \( F = F_{\Gamma} \) in \( \mathbb{H} \) consisting of points \( z = x + yi \) such that \( |z| \geq 1 \) and \(-1/2 \leq x \leq 1/2\). It is known, [160], that \( F \) is a fundamental domain for \( \gamma \) acting on \( \mathbb{H} \).

Suppose \( z_1, z_2 \in F \) and \( \gamma(z_1) = z_2 \) for some \( \gamma \in \Gamma \), then either \( x_1 = -1/2 \) and \( z_2 = T(z_1) = z_1 + 1 \), or \( x_1 = 1/2 \) and \( z_2 = T^{-1}(z_1) = z_1 - 1 \), or \( |z_1| = 1 \) and \( z_2 = S(z_1) = \frac{-1}{z_1} \). Identifying the boundary points of \( F \) which are mapped to each other via elements of \( \Gamma \) we obtain the quotient space \( \Gamma \backslash \mathbb{H} \).

The measure

\[
\frac{dx dy}{y^2}
\]

on \( \mathbb{H} \) is invariant under the action of \( SL_2(R) \uparrow \). It is easy to check that the area of the fundamental domain w.r.t. the invariant measure is finite.

We can give an interpretation of \( F \) (or more precisely \( \Gamma \backslash \mathbb{H} \)) as a space that classifies lattices with homothety as an equivalence relation.
Definition 110. A lattice $L$ in $\mathbb{C}$, the complex numbers, is the $\mathbb{Z}$-span of a pair of $\mathbb{Z}$ independent points $z_1, z_2$, namely all combinations of the form $nz_1 + mz_2$, $m, n \in \mathbb{Z}$, where we assume that non of these combinations, but the trivial one vanishes. We will denote the lattice spanned by $z_1, z_2$ by $\langle z_1, z_2 \rangle$.

We have an action of $\mathbb{C}^* = \mathbb{C} - \{0\}$, the non zero complex numbers, on lattices. An element $z \neq 0$ maps $\langle z_1, z_2 \rangle$ to the lattice $\langle z_1, z_2 \rangle$. Geometrically this amounts to a rotation of the lattice followed by stretching.

Definition 111. If a lattice $L_1$ is mapped to a lattice $L_2$ by the action of $\mathbb{C}^*$, we say that the lattices are homothetic to each other. This is an equivalence relation on lattices.

Given the lattice $\langle z_1, z_2 \rangle$ we may multiply it by $1/z_1$ to get an equivalent lattice $\langle 1, z \rangle$, where $z = z_2/z_1$. Replacing $z$ by $-z$ if needed we may assume that $z \in H$. Essentially by definition, if $\gamma \in SL_2(\mathbb{Z})$, then, the lattice $\langle 1, z \rangle$ is the same as the lattice $\langle az + b, cz + d \rangle$, which up to homothety is the same as the lattice $\langle 1, \gamma(z) \rangle$. Working backwards $\langle 1, z_1 \rangle$ and $\langle 1, z_2 \rangle$ are homothety equivalent lattices if and only if $z_1 = \gamma(z_2)$ for some $\gamma \in SL_2(\mathbb{Z})$ producing our claim that the quotient $\Gamma \backslash H$ classifies lattices up to homothety.

With these preliminaries we can define some special class of functions.

Definition 112. We say that $f$ is weakly modular of weight $k$ w.r.t. $\Gamma = SL_2(\mathbb{Z})$ if it is meromorphic on $H$ and for any $\gamma \in \Gamma$ we have

\begin{equation}
(140) \quad f(z) = (cz + d)^{-k} f(\gamma(z)) = f_{\gamma,k}(z)
\end{equation}

Considering the generators $T, S \in \Gamma$ this is equivalent to

\begin{equation}
(141) \quad f(z) = f(z + 1)
\end{equation}

and

\begin{equation}
(142) \quad f(z) = z^{-k} f(-1/z)
\end{equation}

From the first equation we see that we can think of $f$ as a function of

$$q = e^{2\pi i z}$$

Having $z \in H$ translates into $0 < |q| < 1$, so we have a meromorphic function on this annulus.

Definition 113. If $f$, a weakly modular function, can be extended meromorphically to $q = 0$ (corresponding to "$z = \infty$"), we say that $f$ is a modular function. If $f$ is holomorphic (complex analytic) on $H$ and at infinity $q = 0$, i.e., $f(z) = \sum_{n=0} a_n q^n$ for $|q| < 1$, we say that $f$ is a modular form w.r.t. $SL_2(\mathbb{Z})$ of weight $k$. This power (Fourier) series expansion is known as the $q$-expansion of $f$. We say that $f$ is a cusp form if $a_0 = 0$. 
The name modular form comes from the fact that if \( k \) is even, equations (141) and (142) show that \( f(z)dz^{k/2} \) is invariant w.r.t. \( SL_2(\mathbb{Z}) \), hence, can be defined on \( \Gamma \backslash H \).

Given the interpretation of the quotient as classifying lattices, we can think of a modular form \( g \) as a function \( G \) on lattices \( L \) such that if \( z \neq 0 \) then

\[
G(zL) = z^{-k}G(L)
\]

Conversely, such a function on lattices which is "nice" (holomorphic conditions) will produce a modular form.

Considering the restriction of a modular form \( g(z) \) to the positive imaginary axis given by \( z = it \) and letting \( f(t) = g(it) \), we have

\[
f(t) = g(it) = (-it)^k g(it) = i^k t^k f(1/t)
\]

We see that if \( k \) is a multiple of 4, we get the functional equation \( f = H_{k,1}(f) \).

If the coefficients \( a_n \) of \( g \) are real then the function \( f(t) \) will also be real, since \( q = e^{2\pi iz} = e^{-2\pi it} \) is real on the positive imaginary line. If in addition \( f(t) \) happens to be positive, we will get a nice example of a self-dual generalized density attached to the modular form \( g \). This explains our interest in these forms.

6.2. Modular forms for congruence subgroups. Our quest for more interesting examples leads us to consider modular forms for congruence subgroups.

**Definition 114.** Let \( \Gamma(N) \) be the subgroup of \( SL_2(\mathbb{Z}) \) consisting of matrices \( M \) such that \( M - Id = NM \) for some matrix \( M \), i.e., a matrix which is congruent to the identity modulo \( N \). A subgroup of \( SL_2(\mathbb{Z}) \) which contains the subgroup \( \Gamma(N) \) for some \( N \) is called a congruence subgroup. The congruence subgroup \( \Gamma_0(N) \) consists of those matrices with \( c = 0 \mod N \). The congruence subgroup \( \Gamma_1(N) \) consists of matrices with \( c = 0 \mod N \), \( a = 1 \mod N \) and \( d = 1 \mod N \).

Obviously \( \Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N) \). Let \( \Gamma \) be some congruence subgroup. As in the case of \( \Gamma = \Gamma(1) = SL_2(\mathbb{Z}) \) before we can find a fundamental domain \( F_{\Gamma} \) for the action of \( \Gamma \) on \( H \). We can obtain \( F_{\Gamma} \) from \( F_{\Gamma(1)} \) by taking the union of \( \alpha_i^{-1}(F_{\Gamma(1)}) \), where \( \alpha_i \) provide a set of representatives for the space of cosets \( \gamma \Gamma \), where \( \gamma \in \Gamma(1) \). The set of \( \alpha_i \) is finite since the quotient \( SL_2(\mathbb{Z})/\Gamma(N) = SL_2(\mathbb{Z}_N) \) is finite. By construction we see that the measure of the fundamental domain w.r.t. the invariant measure \( \frac{dx dy}{y^2} \) is finite.

Making the boundary identifications of \( z \) with \( \gamma(z) \), for \( \gamma \in \Gamma \), will "close up" \( F_{\Gamma} \) apart from a few "missing" points on the boundary of \( H \).

**Definition 115.** A cusp of \( F_{\Gamma} \) is a point \( z \in \mathbb{R} \cup \{\infty\} \) where we can find a sequence of points \( z_n \in F_{\Gamma} \) with \( z_n \to z \). Convergence to \( \infty \) being equivalent to \( \text{Im}(z_n) \to \infty \).
Cusps can be either \( z = \infty \) which corresponds after the change of variable \( g = e^{2\pi iz} \) to \( q = 0 \), or rational numbers on the real line. Adding the cusp points to \( F \) will result in a closed surface of a certain genus \( g \) (a bagel/pretzel with \( g \) holes).

We recall that in the case of \( \Gamma = SL_2(\mathbb{Z}) \), the quotient \( \Gamma \backslash H \) had an interpretation as the set of lattices up to homothety equivalence. Similarly, the quotients of \( H \) w.r.t. \( \Gamma_0(N) \) and \( \Gamma_1(N) \) can be interpreted as classifying lattices with extra data, up to homothety. To see this, we first note that a lattice is an additive subgroup of the complex numbers \( \mathbb{C} \). We may consider the compact quotient \( \mathbb{L}_L = \mathbb{C}/L \), where we identify points \( z_1 \) and \( z_2 \) if \( z_1 - z_2 \in L \). Since \( L \) is an additive subgroup, the quotient \( \mathbb{L}_L \) is a commutative group. In \( \mathbb{L}_L \) we may consider the \( N \)-torsion points. These are points \( \tau + L \), such that \( N\tau + L = 0 + L = L \) or equivalently \( N\tau \in L \). For the lattice \( L = \langle z, 1 \rangle \) these elements can be represented using the points \( \tau = \frac{ez + f}{N} \) for \( 0 \leq e, f \leq N - 1 \). As a group, the \( N \)-torsion points are isomorphic to \( \mathbb{Z}_N \times \mathbb{Z}_N \), where \( \mathbb{Z}_N \) denotes the cyclic group on \( N \) elements. As generators we can take \( e_1 = z/N \) and \( e_2 = 1/N \), which are of order \( N \), namely \( Me_i \notin L \) for any \( 0 < M < N \). When \( \gamma \in SL_2(\mathbb{Z}) \) acts on the Lattice \( L \) it maps \( (z/N) + L \) to \( (\frac{ez + f}{N}) + L \) and \( (1/N) + L \) to \( (\frac{e + f}{N}) + L \). We see that the images of \( e_1, e_2 \) depend only on the entries of \( \gamma \) modulo \( N \) and the images will also be of order \( N \). It is also easy to see that any order \( N \) element will be the image of \( e_2 \) under such a transformation. Consider now the set of lattices with the extra data of a given point of order \( N \). By the discussion above, these may be represented by a lattice \( L \) of the form \( \langle z, 1 \rangle \), with the order \( N \) point being \( e_2 \). We also see from the discussion above that an element \( \gamma \in SL_2(\mathbb{Z}) \) will preserve \( e_2 \) in addition to \( L \) iff \( c = 0 \pmod{N} \) and \( d = 1 \pmod{N} \). Since the determinant is 1, this also implies \( a = 1 \pmod{N} \), or in other words \( \gamma \in \Gamma_1(N) \). We conclude that \( \Gamma_1(N) \backslash H \) classifies the data of a lattice and a point of order \( N \). Similarly we may consider the data of a lattice \( L \) together with a subgroup \( G_N \) of \( \mathbb{C}/L \) of size \( N \). Again, we may represent such a pair by \( L \) of the form \( \langle z, 1 \rangle \) together with the subgroup generated by \( 1/N \). The subgroup generated by \( 1/N \) will be mapped to itself if and only if \( c = 0 \pmod{N} \) and \( d \) is invertible modulo \( N \). But these are precisely the conditions for being in \( \Gamma_0(N) \), hence such pairs are classified by the quotient \( \Gamma_0(N) \backslash H \).

We now come to the analogue of a modular form for \( SL_2(\mathbb{Z}) \).

**Definition 116.** Assume that \( \Gamma \) contains \( \Gamma(N) \). We say that a holomorphic function on \( H \) is a **modular form of weight** \( k \) and **level** \( N \) for \( \Gamma \) if it satisfies the weight \( k \) modular functional equation

\[
f_{\gamma, k} = f
\]

for all \( \gamma \in \Gamma \) and in addition, for any \( \delta \in SL_2(\mathbb{Z}) \), the function \( f_{\delta, k}(z) \) has a series expansion \( \sum_{n=0}^{\infty} a_n q_N^n \) at infinity, where \( q_N = e^{2\pi i z/N} \).
We explain the last condition in the definition. Since $\Gamma(N)$ is contained in $\Gamma$ and since $\Gamma(N)$ contains the matrix $T^N$ which is given by

\[
\begin{pmatrix}
1 & N \\
0 & 1
\end{pmatrix}
\]

the modularity condition (144) states that $f(z) = f(z + N)$. which means that $f$ will have a Fourier series expansion in the variable $q_N$.

We make the important observation that if $f$ satisfies the modularity equation (144) for all $\gamma \in \Gamma$ and $\delta$ is some element of $GL_2(\mathbb{Q})$, $\mathbb{Q}$ the rational numbers, then $f_{\delta,k}$ will satisfy the modularity equation for any $\gamma \in \delta^{-1}\Gamma \delta$.

The group $\Gamma(N)$ is a normal subgroup of $\Gamma(1) = SL_2(\mathbb{Z})$, being the kernel of the reduction modulo $N$ homomorphism. We conclude that if $\delta \in SL_2(\mathbb{Z})$ then $f_{\delta}$ will also be modular w.r.t. $\Gamma(N)$ and therefore will also have a Fourier expansion $\sum_{n \in \mathbb{Z}} a_n q_N^n$.

It is easy to check that by using an appropriate $\delta$ we can map any cusp to $\infty$. The condition states that all the resulting $q_N$-expansions are Taylor series expansions, i.e., $a_n = 0$ for $n < 0$, which is a condition of being holomorphic at all cusps.

**Definition 117.** We say that a function $f$ which is a modular form w.r.t. $\Gamma$ is a **cusp form** w.r.t. $\Gamma$ if $a_0 = 0$ in all the expansions above.

Let $e = 0, 1$. Restricting our attention to $\Gamma = \Gamma_e(N)$, let us denote the corresponding spaces of modular forms by $\mathcal{M}_e^k(N)$ and the cusp forms by $\mathcal{S}_e^k(N)$. It is well known that these spaces are finite dimensional, see [62] chapter 3 or [166] chapter 6 for proofs and dimension computations.

The relation between $\hat{f}_{k,N}(z)$ and the queueing theoretic duality map is given, as before, by restricting to the upper imaginary axis $z = it$. Comparing equations (73) and (146), and using the mapping $g(t) = f(it)$ to move...
between functions with a complex variable and functions with a positive real variable, we see that the involutions $H_{k,N}$ and $\hat{f}_{k,N}$ coincide when $z = it$.

6.2.2. Mellin transforms of modular forms. Let us examine the structure of the Mellin transform of a modular form, restricted to the positive imaginary axis, in terms of the $q$-expansion. In fact, while considering the Riemann zeta function as a transform of the theta function, we considered this issue, so our discussion will simply generalize that example. We ignore for the moment issues of convergence.

**Definition 118.** A Dirichlet series is a function of the form

$$\sum_{n \geq 0} a_n n^{-s}$$

If $f = \sum a_n q^n$, the associated Dirichlet series is given by

$$\tilde{L}_f(s) = \sum a_n n^{-s}$$

A simple change of variable from $it$ to $2\pi nt$ shows that

$$L_{q^n}(s) = L_{e^{2\pi i nz}}(s) = (2\pi n)^{-s} L_{e^{-t}}(s) = (2\pi)^{-s} n^{-s} \Gamma(s)$$

Consequently, in terms of the associated Dirichlet series (where it converges), the Mellin transform of a modular form, given in terms of a $q$-expansion is given by

$$L_f(s) = (2\pi)^{-s} \Gamma(s) \tilde{L}_f(s)$$

The convergence of the Mellin transform depends on the asymptotics of $f$ at infinity and $0$. A cusp form decays exponentially fast (like $q^N$) at infinity and at zero, therefore, it’s Mellin transform will converge for all complex values of $s$. As for the convergence of the associated Dirichlet series, we need some estimate on the Fourier coefficients of $f$.

Let $f$ be a cusp form and consider the function

$$\phi(z) = |f(z)| Im(z)^{-k/2}$$

The modularity condition shows that $\phi$ is invariant under $\Gamma$. It is also continuous and hence it will be bounded on a fundamental domain if it is bounded near the cusps. Choosing $\delta \in SL_2(\mathbb{Z})$ which maps a given cusp to infinity we may consider the behavior of $f_{\delta}$ at infinity which by the definition of a cusp form decays exponentially fast as $Im(z) = y \to \infty$ and hence $\phi(z)$ is bounded. we conclude that $|f(z)| \leq cy^{-k/2}$ for some constant $c$. The Fourier coefficient $a_n$ is given by the formula (for $\Gamma_1(N)$),

$$a_n = \int_0^1 f(x + iy) q^{-n}$$

We conclude that

$$|a_n| \leq cy^{-k/2} e^{2\pi ny}$$

and setting $y = 1/n$ we get

$$|a_n| \leq e^{2\pi c n^{k/2}}$$
We see that the associated Dirichlet series converges if \( \Re(s) > (k/2) + 1 \). We see that the Mellin transform of a cusp form (after getting rid of the Gamma function factor) allows us to analytically continue the associated Dirichlet series to the entire complex plane.

6.2.3. The Petersson inner product on modular forms.

**Definition 119.** For \( f, g \in S^k(\Gamma) \) we define

\[
\langle f, g \rangle = \frac{1}{|SL_2(\mathbb{Z}) : \Gamma|} \int_{F_\Gamma} f(z) \overline{g(z)} \frac{dx \, dy}{y^2}
\]

This Hermitian inner product is called the **Petersson inner product**.

The first term \( \frac{1}{|SL_2(\mathbb{Z}) : \Gamma|} \) in the definition of the Petersson inner product, involving the size of the quotient, is a normalizing constant, so that in case \( f, g \) are modular forms w.r.t. different congruence subgroups it will not matter which congruence subgroup we choose for defining the inner product.

The \( \text{Im}(z)^k \) term together with the modularity of \( f, g \) and the use of the invariant metric \( dx \, dy / y^2 \) makes the calculation \( \Gamma \) invariant and independent of the choice of a fundamental domain. By our previous observation, for a cusp form \( f \) the function \( \phi = \phi_f \) which appears in equation (150) is bounded and decays exponentially near the cusps. From this and the finite measure of \( F_\Gamma \) we conclude that if either \( f \) or \( g \) is a cusp form then the inner product is defined (finite).

The Petersson inner product is compatible with the action of elements \( \gamma \in GL_2(\mathbb{Q})^+ \) on modular cusp forms as can be seen from the following lemma whose proof can be found in [62] proposition 5.5.2.

**Lemma 18.** Assume that \( \gamma \) normalizes \( \Gamma \), namely, \( \gamma^{-1}\Gamma\gamma = \Gamma \) and let \( \gamma^* = \text{det}(\gamma)\gamma^{-1} \). Then,

\[
\langle f, g \rangle = \langle f, \gamma^* g \rangle
\]

for \( f, g \in S^k(\Gamma) \).

6.2.4. The diamond operations. The group \( \Gamma_1(N) \) is normal in \( \Gamma_0(N) \). The quotient \( \Gamma_0(N)/\Gamma_1(N) \) may be identified with \( (\mathbb{Z}_N)^* \) by mapping an element \( \gamma \in \Gamma_0(N) \) to \( d \mod N \). We may use this identification to produce an action of \( (\mathbb{Z}_N)^* \) on \( M^k_1(N) \) and \( S^k_1(N) \).

**Definition 120.** For any \( d \in (\mathbb{Z}_N)^* \) pick some \( \gamma \in \Gamma_0(N) \) with bottom right entry equal to \( d \) modulo \( N \), and define the **diamond operator** \( [d](f) = f_\gamma \).

In terms of lattices with an order \( N \) point, such a \( \gamma \) maps \( (L, 1/N) \) to \( (L, d/N) \). By modularity of \( f \) w.r.t. \( \Gamma_1(N) \) the action \([d]\) does not depend on the representative and provides a group action of \( (\mathbb{Z}_N)^* \) which maps cusp forms to themselves. For a linear operator \( T \) on the space of cusp forms, let \( T^* \) denote its adjoint w.r.t. the Petersson inner product, i.e., the operator that satisfies \( \langle Tf, g \rangle = \langle f, T^* g \rangle \). We conclude from (151) that
$[d]^* = [d^{-1}]$, where the inverse is taken in the group $(\mathbb{Z}_N)^*$. In particular these operators are normal, i.e., commute with their adjoints. Since the group $(\mathbb{Z}_N)^*$ is commutative, we have by a basic result in linear algebra a basis of $M^k_1(N)$ consisting of mutual eigenvectors for all the $[d]$. Considering an eigenvector $f$, the eigenvalues w.r.t. $[d]$ will be a group homomorphism of $(\mathbb{Z}_N)^*$ into $C^*$, that we denote by $\chi_f$.

**Definition 121.** Given a homomorphism $\chi : (\mathbb{Z}_N)^* \to C^*$, all the eigenvectors $f$ such that $\chi = \chi_f$ define the $\chi$ **eigenspace** of the diamond operators and are denoted by $M^k_1(N, \chi)$, or $S^k_1(N, \chi)$ for the cusp forms.

The space $M^k_1(N)$ decomposes as a direct sum of these eigenspaces. Consider an element $\gamma \in \Gamma_0(N)$ with lower right element $d$, then it is easy to check that $W_N^{-1}W_N^{-1}$ will have lower right element $a$ with $a = d^{-1} \mod N$. We conclude that if $f \in M^k(N, \chi)$ we have $f_{k,N} = f_{k,N} \in M^k(N, \bar{\chi})$, so the involution is compatible with the decomposition. We also note that by definition $M^k_1(N) = M^k(N, \chi_{\operatorname{triv}})$, where $\chi_{\operatorname{triv}}$ denotes the trivial character.

**6.2.5. Old forms.** We observe that if $K|N$ then $\Gamma_e(N) \subset \Gamma_e(K)$. Therefore, a modular form w.r.t. $\Gamma_e(K)$ is also a modular form w.r.t. $\Gamma_e(N)$. In addition, if $N = LD$, consider the matrix $\delta_D$ given by

$$
(152) \quad \begin{pmatrix} D & 0 \\ 0 & 1 \end{pmatrix}
$$

The matrix $\delta_D$ induces on the upper half plane $H$ the mapping $\delta_D(z) = Dz$. It is easy to verify that

$$
\Gamma_e(N) \subset \delta^{-1}_D \Gamma_e(L) \delta_D
$$

hence for any $f \in M^k_e(L)$ we have

$$
f_{\delta_D,k} = D^{k-1} f(Dz) \in M^k_e(N)
$$

**Definition 122.** A form $f \in M^k_e(N)$ is said to be an **old form** if it is in the subspace of $M^k_e(N)$ which is generated by elements of $M^k_e(K)$ for $K|N$, $K \neq N$, and by elements of the form $f(Dz)$ for $f \in M^k_e(L)$, $LD = N$ and $L \neq N$. A form is called a **newform** if it is orthogonal to all old forms w.r.t. the Petersson inner product.

Old forms, are more likely to be self dual w.r.t. involutions of the form $H_{k,K}$, or $H_{k,L}$, rather than w.r.t. $H_{k,N}$.

We note that the mapping $z \rightarrow Dz$ induces a mapping $q \rightarrow q^D$, which yields a simple calculation of the Fourier expansion of $f_{\delta_D,k}$ in terms of the expansion of $f$.

**6.2.6. Hecke operators.** In addition to the operators $[d]$ and the duality operator $f \mapsto f_{k,N}$ that were noted previously, there is a family of natural operators $T_n$, known as the Hecke operators, acting on $M^k_e(N)$, preserving the cusp forms. This family is of central importance in number theoretic applications of modular forms. For full details we refer the reader to chapter 5 of [62].
We first explain these operators on the level of a lattice $L$ equipped with a point $e_N$ of order $N$. Given some $n$, we can consider all the lattices $L'$, such that $L \subset L'$, $|L'/L| = n$ and $e_N$ still has order $N$ in $L'$. Such lattices $L'$ are obtained from $L$ by adjoining to $L$ all elements in a subgroup $G$ of order $n$ of $E_L = C/L$. If $\gcd(n,N) = 1$, then the point $e_N$ will remain of order $N$ in $L'$, since the subgroup generated by $e_N$ cannot intersect $G$. On the other hand, if $\gcd(n,N) > 1$ then some groups $G$ of order $n$ will intersect the group generated by $e_N$. We exclude the lattices which are obtained from adjoining such groups.

**Definition 123.** Considering a modular form $f$ as a function $F$ evaluated on a lattice $L$ with a point (or subgroup) of order $N$, the $n$th Hecke operator $T_n$ sends it to the modular form $g$ whose corresponding lattice function $G$ satisfies

$$G(L, e_N) = \sum_{L'} F(L', e_N)$$

with $L'$ as described above.

With a bit of elementary, commutative group theory combinatorics, it is not difficult to show from this description that all the operators $T_n$ commute and in fact they also commute with all the operators $[d]$ for $d \in (\mathbb{Z}_N)^*$. It is well known that a finite commutative group of order $nm$ with $\gcd(n,m) = 1$ decomposes as a direct sum of groups of order $n$ and $m$. We can deduce from this fact the multiplicative relation

$$T_{nm} = T_nT_m$$

whenever $n, m$ are prime to each other, namely, $\gcd(n, m) = 1$.

To understand more explicitly how Hecke operators work, let us consider the case $n = p$, a prime number. Consider the lattice generated by 1 and $z$, and let the order $N$ point be $1/N$. In this case the order $p$ points of the lattice form a 2 dimensional vector space over the field $F_p$ with $p$ elements, and subgroups of order $p$ correspond to lines in this vector space. There are $p+1$ lines in the vector space. These lines are generated as subgroups by the points, $(z + j)/p$, for $j = 0, ..., p - 1$ as well as the point $1/p$. Adjoining any of the first $p$ points we obtain a lattice which still has 1 as a generator and the point $1/N$ still has order $N$. Adjoining the point $1/p$, we get a lattice which is generated by $z$ and $1/p$. By multiplication by $p$, we get the lattice generated by 1 and $pz$, while the order $N$ point $1/N$, has been replaced by $p/N$, the same effect that the operator $[p]$ has. However, if $p|N$ there is an intersection of subgroups and we have to discard this lattice altogether from our list. Using this description and some elementary combinatorics of subgroups of $(\mathbb{Z}_{p^e})^2$, we also obtain for $p$ not dividing $N$, the recursive formula

$$T_{p^e} = T_pT_{p^{e-1}} - p^{k-1}[p]T_{p^{e-2}}$$

(154)
for \( r \geq 2 \), where \( T_1 \) is the identity. For \( p|N \) we get
\[
T_p^* = T_p^d
\]
In terms of matrices, the matrices \( \gamma_j \), \( j = 0, ..., p - 1 \), given by
\[
\begin{pmatrix}
1 & j \\
0 & p
\end{pmatrix}
\]
satisfy \( \gamma_j(z) = (z + j)/p \), while the matrix \( \gamma_\infty \), given by
\[
\begin{pmatrix}
p & 0 \\
0 & 1
\end{pmatrix}
\]
satisfies \( \gamma_\infty(z) = pz \). Putting it all together and not forgetting the action of \([p]\) we get
\[
T_p(f) = \sum_{j=0}^{p-1} f_{\gamma_j,k} + [p](f_{\gamma_\infty,k})
\]
the last term being omitted whenever \( p|N \).

We can then easily determine the effect of \( T_p \) on Fourier coefficients, using the identities
\[
q_{\gamma_j,k}^n = p^{k-1}e^{2\pi i n(z+j)/p} = p^{k-1}e^{2\pi i (n/p)z}e^{2\pi ij/p}
\]
and
\[
q_{\gamma_\infty,k}^n = p^{k-1}q^n
\]
Using the elementary fact (see chapter 4) that \( \sum_{j=0}^{p-1} e^{2\pi i nj/p} = 0 \), if \( p, N \) are prime to each other, and equals \( p \) otherwise we can show that in terms of the \( q \) expansion
\[
T_n(f) = \sum_m \left( \sum_{d|gcd(m,n)} d^{k-1}a_{mn/d^2}([d]f) \right) q^m
\]
For \( f \in M_k^1(\chi) \) this reduces to
\[
T_n(f) = \sum_m \left( \sum_{d|gcd(m,n)} d^{k-1}a_{mn/d^2}(f)\chi(d) \right) q^m
\]
applying this equation with \( m = 1 \) we see that for any \( n \) and \( f \in M_k^1(\chi) \) we have
\[
a_1(T_n,f) = a_n(f)
\]
The Hecke operators relate nicely to the operators \( W_N \), in fact, it can be shown, \([55]\), that for all \( n \),
\[
W_N T_n W_N^{-1} = T_n^*
\]
It can also be shown that if \( gcd(p,N) = 1 \) then
\[
T_p^* = [p]^{-1}T_p
\]
In particular, these Hecke operators are normal. We conclude that the set of operators consisting of \([m]\) and \( T_m \) for all \( m \) prime to \( N \) is a set of normal
commuting operators acting on the finite dimensional vector spaces $S^k_e(N)$ and $M^k_e(N)$. As such they have a basis of mutual eigenvectors. Consider a mutual eigenvector $f = \sum_n a_n(f)q^n$. Let $n$ be prime to $N$ and denote the corresponding eigenvalue of $T_n$ on $f$ by $\lambda_n$. Comparing with (160) we obtain $a_n(f) = a_1(f)\lambda_n$.

**Definition 124.** We say that a nonzero modular form is a **Hecke eigenform** if it is an eigenvector for the operators $[n], T_n$ for all $n$.

By definition, the space of old forms is preserved by the Hecke operators $T_n$, the diamond operators $[n]$, for all $n$, and by $W_N$. Using (161) it is also preserved by the adjoints of the Hecke operators, thus the space of newforms, being orthogonal to the space of oldforms, is also stable under the action of all Hecke operators. The following is a theorem of Atkin and Lehner, [12], see [62] for a proof.

**Theorem 59.** If $f = \sum a_nq^n \in S^k(N)$ and for all $n$ such that $n, N$ are prime to each other we have $a_n = 0$ then $f$ is an old form.

As a consequence we prove the following result.

**Theorem 60.** Let $f$ be a newform which is an eigenvector w.r.t. all $T_n$ and $[n]$ with $\gcd(n, N) = 1$. Then, $f$ is an eigenvector w.r.t. $T_n$ and $[n]$ for all $n$.

**Proof:** Consider a newform $f$ which eigenvector $f$ w.r.t. all $T_n$ and $[n]$ such that $\gcd(n, N) = 1$. Assume that $a_1 = 0$. By (160) we have $a_n = 0$ for all $n$ such that $\gcd(n, N) = 1$, and by theorem 59 we find that $f$ is an old form and we obtain a contradiction. Consequently, we conclude that $a_1 \neq 0$. Given any $m$, the function $g = T_m f - a_m f$ will again be an eigenform for all $T_n$ and $[n]$ with $\gcd(n, N) = 1$, because all Hecke operators commute. We also have $a_1(g) = 0$, therefore $g$ is an old form. But, $g$ is also a new-form since Hecke operators preserve newforms, hence $g = 0$ and we conclude that $f$ is an eigenvector for $[n]$ and $T_n$ for all $n$. q.e.d.

**Definition 125.** A newform which is a cusp form and is an eigenvector for the operators $T_n$ and $[n]$ for all $n$, and which is normalized so that $a_1 = 1$ is called a **Hecke newform**.

The Hecke newforms form an orthogonal basis for the space $S^k_1(N)^{new}$ of new cusp forms. For a Hecke newform, $a_n$ is the eigenvalue corresponding to the operator $T_n$.

**Theorem 61.** Let $f$ be a Hecke newform, and let $g(t) = f(it)$, then either, $g = H_{k,N}(g)$ or $g = -H_{k,N}(g)$.

**Proof:** Consider the subspace of newforms $S^k_0(N)^{new} \subset S^k_0(N)$. From equations (161) and (162) we see that if $f \in S^k_0(N)^{new}$ is a Hecke eigenform, then $f_{W_N,k}$ is also a Hecke eigenform and they have the same eigenvalues.
w.r.t. $T_n$, such that $gcd(n, N) = 1$. The vanishing $a_1$ trick shows that any two Hecke newforms with the same eigenvalues w.r.t. all the operators $T_n$ with $gcd(n, N) = 1$ are equal. We conclude that $f$ and $f_{W_k,N}$ are proportional to each other and so $\hat{f}_{k,N} = f$ or $\hat{f}_{k,N} = -f$. Restricting to the upper imaginary axis we obtain our result. q.e.d.

The Dirichlet series associated with a Hecke eigenform of $\Gamma_0(N)$ is characterized by a property which is of great importance in number theoretic applications, it has a particular product formula. It is easy to check that the equations (153,154,155) are equivalent to the statement that the associated Dirichlet series satisfies

\begin{equation}
\tilde{L}_f(s) = \prod_p (1 - a_p p^{-s} + 1_N(p)p^{k-1-2s})^{-1}
\end{equation}

where $1_N(p) = 0$ if $p$ divides $N$ and is equal to 1 otherwise.

7. Hecke newforms and the mini-market theorem

We want to relate Hecke newforms of weight $k = 2$ to the functions $g_{A,B}$ from the mini-market theorem. We first want to explain the meaning of the conditions that we imposed on the integers $A$ and $B$. Recall that we asked that $(4A^3 + 27B^2) \neq 0$ and that if $p^4$ divides $A$ then $p^6$ does not divide $B$.

We are interested in the set of solutions of the equation

$P(x, y) = P_{A,B}(x, y) = y^2 - (x^3 + Ax + B) = 0$

If we consider this equation with the variables $x, y$ being complex numbers, the set of solutions is a subset of $\mathbb{C}^2$. It can be shown that it will be a smooth manifold in the sense of chapter 3 if and only if at each solution point $(x, y)$ the gradient vector satisfies

$$
\left( \frac{\partial P}{\partial x}, \frac{\partial P}{\partial y} \right) = (2y, -3x^2 - A) \neq 0
$$

We see that vanishing of the gradient requires $y = 0$. Since we are only considering pairs $(x, y)$ which are solutions to $P_{A,B}(x, y) = 0$ we conclude that vanishing of the gradient can occur only at roots of the cubic $Q(x) = Q_{A,B}(x) = x^3 + Ax + B$ and only if the derivative $Q'(x)$ vanishes at the root. The derivative of a polynomial $Q$ will vanish at a root only if it is a double or higher multiplicity root. Thus the gradient never vanishes if and only if the three complex roots of the cubic $Q(x)$ are distinct.

If the roots of $Q$ are $e_1, e_2, e_3$, then equivalently we may ask that the discriminant $\Delta$ of the cubic which is given by $\Delta = (e_1 - e_2)^2(e_2 - e_3)^2(e_1 - e_3)^2$ be non zero. In terms of the coefficients $A$ and $B$, the discriminant of $Q_{A,B}$ is given by

$$
\Delta = -(4A^3 + 27B^2)
$$

We see that the geometric interpretation of the condition $4A^3 + 27B^2 \neq 0$ in the mini-market theorem is that the set of complex solutions is a (nice)
smooth manifold. Note that we never used in the discussion the fact that $A$ and $B$ are integers, they could be any elements in $C$.

The entire discussion above can also be carried out without change if we replace the complex numbers $C$ by any algebraically closed field $K$. While we do not have a topology as in the case of the complex numbers we can still use the non-vanishing gradient condition as an algebraic analogue for the notion of a smooth manifold. Fields of characteristic 2 or 3 are rather problematic and we avoid these primes throughout this subsection. We are led to the following definition.

**Definition 126.** If the discriminant of $Q_{A,B}$ is non zero we say that the set of solutions of $P_{A,B}(x, y) = 0$ is non-singular. If the set of solutions is non-singular then we will also say that the set of solutions forms an elliptic curve which we denote by $E_{A,B}$.

We now consider the equation $P(x, y) = 0$ as a polynomial equation with integer coefficients. We can make the linear change of variables $(x, y) \mapsto (c^{-2}x, c^{-3}y)$, with $c$ a rational number. This change of variables provides a linear bijection between the set of solutions to our original equation $P(x, y) = y^2 - (x^3 + Ax + B) = 0$ and the set of solutions of the equation $P_c(x, y) = y^2 - (x^3 + c^{-4}Ax + c^{-6}B) = 0$, in any field $F$ which contains the rational numbers, i.e., any field of characteristic 0. If $c = \frac{n}{m}$ where $n, m$ are prime to each other, then the new equation $P_c(x, y)$ will have integer coefficients if and only if for each prime $p$ dividing the numerator $n$ we have $p^4 | A$ and $p^6 | B$. The condition in the mini-market theorem on the divisibility properties of $A, B$ guarantees that the numerator can only have values $n = 1, -1$. The change of variable changes $\Delta$ to $c^{-12}\Delta = m^{12}\Delta$, the later equality when $n = 1, -1$.

**Definition 127.** An equation $P(x, y)$ with integer coefficients whose associated discriminant divides the discriminant of any equation $P_c(x, y)$ obtained from it using a change of variables as above is called a minimal discriminant equation.

We conclude that under the conditions of the mini-market theorem, the equation will be a minimal discriminant equation.

Given that the coefficients of $P$ are integers, we may reduce the coefficients of $P(x, y)$ modulo $p$ and obtain an equation $P_p(x, y)$ with coefficients in $\mathbb{F}_p$, the field with $p$ elements. Since the discriminant is given by a polynomial expression in the coefficients, the discriminant of $P_p$ will be the reduction modulo $p$ of the discriminant of $P$. If $p$ does not divide the discriminant, it will remain non-zero after the reduction and we will obtain a non-singular curve.

**Definition 128.** Consider an elliptic curve $E$ given by an equation $P(x, y) = 0$ with integer coefficients. Let $p$ be a prime such that $p$ does not divide the discriminant $\Delta_E$, then we say that $E$ has good reduction
at the prime $p$ or that $p$ is a prime of good reduction. We say that $E$ has **bad reduction** at $p$ if it does not have good reduction.

Primes of bad reduction further split into two cases. If only two of the three solutions to the cubic $Q_p$ in the algebraic closure of $F_p$ coincide, we say that $E$ has **multiplicative reduction**. Otherwise, if all three solutions coincide we say that $E$ has **additive reduction**.

We recall that $g_{A,B} = \sum_n a_n q^n$ where $a_{p^m} = p^m - N_{p^m}$ and $N_{p^m}$ is the number of solutions of $P_{p^m}$ in the field $F_{p^m}$ with $p^m$ elements.

Let us consider the computation of $a_{p^m}$ in the case of bad reduction. A typical equation for a curve of bad reduction is of the form

$$y^2 = x^3 + Ax^2$$

with $A \in F_p$. This is the case when $x = 0$ is a multiple root of the cubic $Q(x)$. The case of additive bad reduction, with a triple root corresponds to $A = 0$, while $A \neq 0$ corresponds to a double root or multiplicative reduction. Consider the variable $t = \frac{y}{x}$. If $x \neq 0$ we may rewrite the equation (after dividing by $x^2$) as

$$t^2 = x + A$$

One solution is $(0, 0)$. Any choice of $t \in F_{p^m}$ will lead to another solution, unless $t^2 = A$ in which case we get $x = 0$ which means that the modified equation is not valid. The solutions for different values of $t$ will be different since $t$ is determined by $x, y$.

If $A = 0$, the additive reduction case, we get a single illegal choice for $t$, namely $t = 0$. The remaining $p^m - 1$ choices give $p^m - 1$ solutions, adding to them the point $(0, 0)$ we get a total of $p^m$ solutions. Applying our formula for computing $a_{p^m}$ we get $a_{p^m} = 0$.

If $A \neq 0$ is a quadratic residue in $F_p$ then there are 2 illegal values for $t$ corresponding to the square roots of $A$, we thus get $p^m - 2$ solutions and adding the point $(0, 0)$ we get $p^m - 1$ solutions and so $a_p = 1$. We remark that this case is known as the **split multiplicative reduction** case.

If $A$ is not a quadratic residue then any of the $p$ values of $t$ is legal and we get $a_p = -1$. On the other hand, we can adjoin to $F_p$ the root of $A$. This will lead to a quadratic extension of $F_p$, i.e., a field which is a two dimensional vector space over $F_p$ with basis $1, \sqrt{A}$. The unique up to isomorphism quadratic extension of $F_p$ is $F_{p^2}$, hence that field contains the root of $A$. Therefore, as in the split case $a_{p^2} = 1$. The field with $p^m$ elements consists of the solutions to the equation $x^{p^m} = x$ in the algebraic closure of $F_p$. For $m$ even, the field with $p^m$ elements, thus contains the field with $p^2$ elements, so for these fields $A$ is quadratic and $a_{p^m} = 1$. When $m$ is odd, then $F_{p^m} \cap F_{p^2} = F_p$ since $F_{p^k}$ is not a vector space over $F_{p^2}$ and so these fields do not contain a root of $A$. We conclude that $a_{p^m} = (-1)^m$. This case is known as the **non-split multiplicative case**.

In the case of good reduction, we have the following important theorem due to Hasse, [92], see [62] for a proof.
Theorem 62. Let $E$ be an elliptic curve with good reduction at $p$, then $a_{p^m} = \lambda_p^m + \bar{\lambda}_p^m$, where $\lambda_p \in \mathbb{C}$ satisfies $|\lambda_p|^2 = \lambda_p \bar{\lambda}_p = p$.

The theorem leads to the recurrence

$$a_{p^m} = \lambda_p^m + \bar{\lambda}_p^m$$

$$= (\lambda_p^{m-1} + \bar{\lambda}_p^{m-1})(\lambda_p + \bar{\lambda}_p) - (\lambda_p \bar{\lambda}_p)(\lambda_p^{m-2} + \bar{\lambda}_p^{m-2})$$

$$= a_{p^{m-1}}a_p - pa_{p^{m-2}}$$

where we have assumed $a_1 = 1$ and $m \geq 2$.

Putting it all together, and recalling that by definition $a_{nm} = a_n a_m$ whenever $\gcd(n, m) = 1$, we see that the $a_n$ associated with an elliptic curve $E$, satisfy the same relations as the Fourier coefficients of a Hecke eigenform in $S_2^0(N)$ where $N$ is divisible by the same primes as the discriminant $\Delta_E$.

We are ready to consider some wishful thinking and ask whether the function $g_E = \sum_n a_n q^n$ happens to be a Hecke eigenform and if so, of what level?

Definition 129. Let $E$ be an elliptic curve defined by an equation $y^2 = Q(x)$ with integer coefficients. We define $N_E$, the conductor of $E$, as follows. Let $p \neq 2, 3$ be a prime. Define $N_{p,E} = 1$, if $p$ has good reduction, $N_{p,E} = p$ if $p$ has multiplicative reduction and $N_{p,E} = p^2$ if $p$ has additive reduction. An explicit computation of $N_{p,E}$ when $p = 2, 3$ uses a formula due to Ogg, [139] and an algorithm due to Tate, [170]. Both can be found in [162]. We let $N_E = \prod p, N_{p,E}$.

We note that $N_E$ and $\Delta_E$ are divisible by the same primes.

Definition 130. An elliptic curve $E$ is said to be modular if $g_E = \sum_n a_n q^n$ is the $q$-expansion of a cusp form.

The following result is known as the Taniyama-Shimura conjecture. It was established in [41], relying heavily on the work of [186] and [171], which established an important special case. The proof is extremely complicated and is considered as one of the highlights of mathematics. For an overview we refer the reader to [55].

Theorem 63. Let $E$ be an elliptic curve with minimal discriminant, then $g_E$ is a Hecke newform in $S_2^0(N_E)$.

We can now prove the mini-market theorem.

Proof of theorem 53: By theorem 63, the function $g_{A,B}$ is a Hecke newform of weight $k = 2$ and level $N_E$. This implies exponential decay at 0 and $\infty$, hence the same holds for $\tilde{f} = g_{A,B}^2$. We conclude that $\int_0^\infty \tilde{f}(t)dt < \infty$ proving claim 0) of the theorem. In addition theorem 61 tells us that $H_{2,N_E}(g_{A,B}) = g_{A,B}$ or $H_{2,N_E}(g_{A,B}) = -g_{A,B}$. This implies that $H_{4,N_E}(f) = f$ and so by theorem 57, the other claims of the theorem are established. q.e.d.
Remark: One major motivation for proving that $g_{A,B}$ is a Hecke newform was proving the analytical continuation and functional equation for the function $L_{g_{A,B}}$, as was the case with Riemann’s zeta function. However, another major motivation (in fact, the main motivation for Wiles) is that together with results of Frey, Serre and Ribet it resolved Fermat’s last theorem. See [55] for details.

8. SITA queues with Pareto job size distributions

We now discuss the behavior of SITA queues in a case that is relatively easy to analyze, when the job size distribution is Bounded Pareto. We will see that the behavior fits well with duality and that the Pareto distribution with parameter $\alpha = 1$ has some extremal queueing theoretic properties among all distributions. Since SITA policies work by reduction of variance we are especially interested in the cases where the coefficient of variation $C^2_\alpha = E(X^2)/E(X)^2$ of the Pareto distribution is large. A short computation shows that for $0 < \alpha < 2$ the coefficient of variation is infinite on the interval $[1, \infty)$, hence, also the normalized average waiting time. We consider this case by looking at the Bounded Pareto version $P_{\alpha,l}$ with generalized density $f = f_{\alpha,l}$, for fixed $\alpha$ and consider the behavior of the queue as $l$, the ratio of longest to shortest job tends to infinity. We also fix the number of servers in the queueing system $h$ and assume that they are all identical. Finally, we fix the load $\rho$. This setting has been a standard testing ground for SITA in several papers, [158, 85, 87]. Let us first consider the case where the load is low $\rho < 1$. This means that a single host can handle all the load in the system and so we can choose the cutoffs $1 = s_0(l) < s_1(l), ..., s_{h-1}(l) < s_h(l) = l$ freely without the fear of overloading any host. By duality, the weight 4 dual of $P_{\alpha,l}$ is $P_{2-\alpha,l}$, hence, it is enough to consider $2 > \alpha \geq 1$. We first assume $\alpha > 1$, the formulas for $\alpha = 1$ are slightly different. We sometimes will abbreviate $s_j(l)$ to $s_j$. If we have functions $f(l), g(l)$ we will write $f \sim g$, if $\lim_{l \to \infty} f(l)/g(l) = 1$.

We compute the normalized average waiting time using equations (101, 102).

\begin{align*}
L_f(i+1) & = \int_1^l s^i s^{-\alpha} ds = \frac{1}{\alpha - i} \left(1 - \left(\frac{1}{l}\right)^{\alpha - i}\right) \tag{164}
\end{align*}

and similarly

\begin{align*}
L_{f,I_j}(i+1) & = \int_{s_{j-1}}^{s_j} s^i s^{-\alpha} ds = \frac{1}{\alpha - i} \left(s_j^{\alpha + i} - s_{j-1}^{\alpha + i}\right) \tag{165}
\end{align*}

which leads to

\begin{align*}
\frac{L_{f,I_j}(1)L_{f,I_j}(3)}{(L_f(2))^2} & = \frac{(1-\alpha)^2}{\alpha(2-\alpha)} s_{j-1}^{-\alpha} s_j^{-2\alpha} (1 - \left(\frac{s_{j-1}}{s_j}\right)^\alpha) (1 - \left(\frac{s_{j-1}}{s_j}\right)^{2-\alpha}) \frac{1}{(1 - \left(\frac{1}{l}\right)^{\alpha - 1})^2}
\end{align*}
Using scale invariance it is not difficult to see that as \( l \to \infty \), in any reasonably good set of cutoffs, the successive ratios \( s_j(l)/s_{j-1}(l) \to \infty \), since otherwise the service capacity of the \( j \)'th host is waisted. Making this assumption we see that

\[
L_{f,I_j}(1)L_{f,I_j}(3) \sim (L_{f}(2))^2 \sum_{j=1}^{\alpha} g_j s_{j-1}(l)^{-\alpha} s_j(l)^{2-\alpha}
\]

For the other moment expression in \( E_j(N) \) we have

\[
\frac{L_{f,I_j}(2)}{L_{f}(2)} = (s_j^{\alpha+1} - s_{j-1}^{\alpha+1})(1 - (\frac{1}{l})^{\alpha-1})^{-1}
\]

For \( j > 1 \) this gives

\[
\frac{\rho}{2(1 - \rho)} \sim \frac{\rho}{2(l)}
\]

and for \( j = 1 \) we have

\[
\frac{\rho}{2(1 - \rho)} \sim \frac{\rho}{2(l)}
\]

We conclude that if for all \( j \), \( s_j(l)/s_{j-1}(l) \to \infty \) we get a very simple asymptotic formula for the average normalized waiting time

\[
E(N)(\alpha, h, \rho, f_\alpha, l, s_1(l), \ldots, s_{h-1}(l)) 
\]

\[
\sim \sum_{j=1}^{h} g_j s_{j-1}(l)^{-\alpha} s_j(l)^{2-\alpha} = G_{\alpha, \rho}(s_1(l), \ldots, s_{h-1}(l))
\]

where

\[
g_1(\alpha, \rho) = \frac{\rho}{2(1 - \rho)} \frac{(\alpha - 1)^2}{(2 - \alpha)\alpha}
\]

and

\[
g_j(\alpha, \rho) = \frac{\rho}{2(1 - \rho)} \frac{(\alpha - 1)^2}{(2 - \alpha)\alpha}
\]

for \( j > 1 \).

The function \( G_{\alpha, \rho} \) has a (unique) minimum over the simplex \( \Delta_{h,l} \) of all possible cutoffs for the job size distribution with density \( f_\alpha, l \). The minimum is attained at a set of cutoffs \( s_1^{\min}(l), \ldots, s_{h-1}^{\min}(l) \). Differentiating \( G \) w.r.t. the \( j \)'th variable and equating to zero leads to the recursive relation

\[
\frac{s_j^{\min}(l)}{s_{j+1}^{\min}(l)} = \frac{2 - \alpha}{\alpha} \frac{g_j}{g_{j+1}} \frac{s_j^{\min}(l)}{s_{j-1}^{\min}(l)} \frac{1}{2-\alpha} \frac{1}{s_j^{\min}(l)}
\]

Together with the telescopic product

\[
\Pi_j(l)(s_j^{\min}(l)/s_{j-1}^{\min}(l)) = s_h^{\min}(l)/s_0^{\min}(l) = l/1 = l
\]
we can solve for $s_j^{\min}(l)$ and the minimal value $G_{\alpha,h,\rho}(s_1^{\min}(l),...,s_{h-1}^{\min}(l))$. Putting it all together the final result is

**Theorem 64.** Assume $\alpha > 1$ and $\rho < 1$. Let $q = \frac{\alpha}{2-\alpha} > 1$. Let $E(N)^{opt}_{\rho,h,\alpha}(l)$ denote the minimal value of $E(N)(\rho, h, f_{\alpha,l}, s_1, ..., s_{h-1})$ over all possible cutoffs $1 < s_1, ..., s_{h-1} < l$, then

$$E(N)^{opt}_{\rho,h,\alpha}(l) \sim (1 - q^{-h})q^{\frac{1}{q-1}} \frac{h}{q^{h-1}} \left( \frac{1}{1-\rho} \right) \left( \frac{q-1}{q^{h}} \right) \rho \left( \frac{\alpha - 1}{4(2-\alpha)} \right) \frac{2\alpha - 2}{l^{2\alpha-2}}$$

We see that $E(N)^{opt}_{\rho,h,\alpha}(l)$ has the form $c(\alpha, h, \rho) l^{\frac{2\alpha - 2}{\alpha - 1}}$. As $l \to \infty$, $E(N)^{opt}_{\rho,h,\alpha}(l) \to \infty$. This is not surprising since one of the intervals $I_j$ must be unboundedly long and for $0 < \alpha < 2$ the single server average normalized waiting time tends to infinity. Recalling that $q = \frac{\alpha}{2-\alpha} > 1$ we see that the rate in which the average normalized waiting time tends to infinity drops dramatically as the number of servers increases as the exponent of $l$ decays exponentially fast. For example, if $\alpha = 3/2$, $\rho = 1/2$ and $h = 5$ we get $c(\alpha, h, \rho) < 3/4$, $q = 3$, and $l^{\frac{2\alpha - 2}{\alpha - 1}} = l^{1/242}$. We see that $l$ has to be ridiculously big, to make this estimate large.

The sore spot for SITA queues is $\alpha = 1$. For this case we can use duality theory to compute the optimal cutoffs precisely.

Since the Pareto distribution with $\alpha = 1$ is self dual with weight 4, if $h = 2$, the optimal cutoff is $s_1 = l^{1/2}$, the fixed point of $t \to l/t$, which maps the interval $[1,l]$ to itself. If we have $h$ servers and optimal cutoffs, then each two server subsystem, spanning jobs in the interval $[s_{j-1}, s_{j+1}]$, must have $s_j$ as optimal cutoff. The scale invariance of the Pareto distribution says that the restriction to $[s_{j-1}, s_{j+1}]$ is self dual with weight 4 and hence $s_j = \sqrt{s_{j-1}s_{j+1}}$ as the fixed point of the appropriate involution. These conditions can only be satisfied if $s_j = l^{1/h}$, so these are the optimal cutoffs. Using the same arguments (duality+scale invariance) we see that all single host average waiting times are equal and that all single server loads are equal to $\rho/h$. These assertions hold without the restrictive assumption $\rho < 1$, so the calculation works for general loads. A direct calculation with the optimal $s_j$ now yields

$$E(N)^{opt}_{\rho,h,1}(l) = \frac{\rho}{2(1-\rho/h)} h(l^{1/h} - 1)^2 \frac{h^{1/h}}{1/h \log^2(l)} \sim \frac{\rho}{2(1-\rho/h)} h^{1/h} \log^2(l)$$

Ignoring the term $\log^2(l)$ which is polynomial in the logarithm of $l$, we see that the average normalized waiting time behaves like $l^{1/h}$, namely, the exponent of $l$ decays only polynomially fast. We show that in this respect, at low loads, $\rho < 1$, the Pareto distribution with $\alpha = 1$ displays the worst possible asymptotic behavior among all distributions.
Theorem 65. Let $X$ be any job size distribution for which $\Pr(X < 1) = \Pr(X > l) = 0$ let $\rho < 1$ be the load of a SITA system with $h$ identical hosts. Let $s_j(l) = \bar{b}^j/h$, then

$$E(N) = E(W)/E(X) \leq \frac{\rho}{2(1-\rho)} \frac{(l^{1/h} + 1)^2}{4l^{1/h}}$$

Proof: Consider server $j$ and let $X_j$ denote the restriction of $X$ to the interval $I_j$. By our assumption that $\rho < 1$, none of the servers will be overloaded. By the P-K formula we have

$$E(W_j)/E(X_j) = \rho_j E(X_j^2)/E(X_j)^2/2(1-\rho_j)$$

By the choice of $s_j$, $X_j$ is supported on the interval $[l^{(j-1)/h}, l^{j/h}]$. Since $E(X_j^2)/E(X_j)^2$ is time scaling invariant we can rescale the time unit, dividing it by $l^{(j-1)/h}$ to obtain a distribution in the range $[1,l^{1/h}]$ with the same normalized waiting time. Let $r = l^{1/h}$.

Next, we note that $E(N) = E(W)/E(X) = \sum_j \rho_j E(X_j)/(E(W_j)/E(X_j))$, where $\rho_j$ is the probability that a job will be serviced by host $j$. This shows that $E(N)$ is a weighted average of $E(W_j)/E(X_j)$, hence it is enough to prove the bound for each $j$ separately.

As explained above, we assume that $X_j$ is supported on the interval $[1, r]$ where $r = l^{1/h}$. Consider a distribution $Z$ which is supported on the same interval and choose some $\delta > 0$. Divide the interval $[1+\delta, r-\delta]$ into a finite number of sub-intervals of length at most $\delta/2$. Let $[a, b]$ be one of the sub-intervals and let $v = \min(a - 1 - \delta/2, r - b - \delta/2)$. We shift half the measure $Z$ restricted to $[a, b]$, a distance $v$ to the right and the other half of the measure, a distance $v$ to the left. Let $\tilde{Z}$ be the resulting measure. We notice that if $v = a - 1 - \delta/2$ then the half which is shifted to the left ends up in the interval $[1 + \delta/2, 1 + \delta/2 + (b-a)]$ which by assumption is contained in $[1, 1 + \delta]$. Likewise, if $v = r - b - \delta/2$ then the half which is shifted to the right ends up in $[r - \delta, r]$. The average remains the same $E(Z) = E(\tilde{Z})$ since we shifted equally the same measure to the left and to the right. However, if $Z$ has positive measure on the interval $[1 + \delta, r - \delta]$, $E(\tilde{Z}^2)$ has strictly increased since $(t + v)^2 + (t - v)^2 = 2t^2 + 2v^2 > 2t^2$. Applying the shifting process to all sub-intervals, we obtain a distribution $Z_1$ for which $E(Z_1^2)/E(Z_1)^2 \geq E(Z^2)/E(Z)^2$ and such that the mass (measure) of $U$ in the subinterval $[1 + \delta, r - \delta/2]$ is at most, half the mass of $Z$ on the same subinterval. Repeating inductively the process starting with $Z_1$, we obtain a distribution $Z_m$ such that $E(Z_m^2)/E(Z_m)^2 \geq E(Z^2)/E(Z)^2$ and the mass of $Z_m$ on $[1 + \delta, r - \delta/2]$ is at most $1/2^{m-1}$. Choosing $m$ large enough we see that there is a distribution $Z_m$ such that $E(Z_m^2)/E(Z_m)^2 \geq E(Z^2)/E(Z)^2$ and that at least $1 - \delta$ of the mass of $Z_m$ is concentrated on the intervals $[1, 1 + \delta]$ and $[r - \delta, r]$. We conclude that the supremum of $E(X^2)/E(X)^2$ over all distributions supported on $[1, r]$ is the same as the supremum over distributions whose mass in the union of the intervals $[1, 1 + \delta]$ and $[r - \delta, r]$.
is at least $1 - \delta$. We take a sequence $\delta_n \to 0$ and distributions $Y_n$ as above corresponding to $\delta_n$, with $E(Y_n^2)/E(Y_n)^2 \to M$ where $M$ is the supremum of $E(Z^2)/E(Z)^2$. We note that the set of distributions which are supported at the points 1 and $r$ is compact, and consists of distributions $Y_w$ which place weight $w$ at the point 1 and weight $1 - w$ at the point $r$. Let $w(Y_n)$ be the measure of the interval $[1, 1 + \delta]$ w.r.t. $Y_n$. Let $w$ be the limit of a sub-sequence of $w(Y_n)$. It is easy to see that $E(Y_w^2)/E(Y_w)^2 = M$, hence the supremum is attained. We can now compute the maximizing $Y_w$ by differentiating $E(Y_w^2)/E(Y_w)^2$ w.r.t. $w$. It turns out that the maximizing $Y_w$ is the unique $H_{4,1/r}$ self-dual distribution which is supported on the points 1 and $r$, namely $w = \frac{r}{r+1}$. A short computation reveals that $E(Y_w^2)/E(Y_w)^2 = \left(\frac{r+1}{4r}\right)^2$ and substituting $l^{1/h}$ for $r$ we obtain the desired result. \emph{q.e.d.}

In the proof we have shown that for a single server the job size distribution that leads to the worst behavior in an M/G/1 queue is the unique $H_{4,1}$ self-dual distribution which is supported on the points 1 and $l$. A similar argument can be used to show that for $h$ servers, the job size distribution, supported on the interval $[1, l]$, which leads to the worst possible average waiting time for an $h$ server SITA queue, is the unique distribution supported on the points $l^{j/h}$, $j = 0, ..., h$, whose restriction to any pair of points $l^{j/h}, l^{(j+1)/h}$ is self-dual w.r.t. $H_{4,(j+1)/h}$. The limit of these distributions $X_h$ is the distribution $P_{1,l}$. We see that for average waiting time, $\alpha = 1$ is the worst parameter for a Pareto distribution and asymptotically is among the worst distributions over all. What about slowdown, where the self-dual distribution has parameter $\alpha = 1/2$. Following the same line of computations that led to theorem 64 we get the following results for average slowdown. We will restrict ourselves to order of magnitude computations, although the constants involved are easily determined as well. By duality it is enough to consider the case $\alpha \geq 1/2$.

\textbf{Theorem 66.} Assume $\rho < 1$. Let $\alpha > 1$, then,

$$E(S)^{\text{opt}} \sim c_{h,\rho,\alpha}(l^{\frac{2\alpha - 1}{\alpha - 1}})$$

where $\hat{q} = \frac{\alpha + 1}{2\alpha}$.

For $1/2 < \alpha < 1$

$$E(S)^{\text{opt}} \sim c_{h,\rho,\alpha}(l^{\frac{2\alpha - 1}{\alpha - 1} l^{\alpha - 1}})$$

In particular, if the number of hosts is large enough, then

$$E(S)^{\text{opt}}_{h,\rho,\alpha}(l) \to 0$$

The same conclusion holds for $\alpha = 1/2$, which offers the best asymptotic performance. The worst asymptotic performance occurs at $\alpha = 1$, for any number of hosts $h$. 

The tension between the weight 3 and weight 4 dualities which are related to slowdown and normalized waiting time respectively, can be seen when we consider what happens to the average slowdown when we optimize the \( s_j \) to give optimal average normalized waiting time and vice versa.

Assume first that the parameters \( s_j \) are chosen to optimize average normalized waiting time. It is easy to show using equation (170) that for the optimal \( s_j \) the terms \( E_j(N) \) will all have the same order of magnitude, namely, their ratios will be bounded. We have

\[
E_j(S) = \frac{L_f(2)}{L_f(1)} E_j(N) L_{f,I_j}(0)/L_{f,I_j}(1)
\]

Whenever, \( s_j/s_{j-1} \to \infty \), then \( L_{f,I_j}(0)/L_{f,I_j}(1) \) has an order of magnitude which is larger than that of \( L_{f,I_j}(0)/L_{f,I_j}(1) \), for \( j > 1 \). We conclude that only \( E_1(S) \) contributes asymptotically to \( E(S) \). We get that for \( \alpha > 1 \)

\[
E(S) \sim c(s_1^{2-\alpha}(l)) = dl^{2\alpha-2}/q^{\alpha-1}
\]

For some constants \( c, d \). This means that the average slowdown is well behaved, with rapid decay as \( h \) increases.

For \( \alpha < 1 \) the same argument yields

\[
E(S) \sim c(l^{1-\alpha/q^h+\alpha-1})
\]

where in this case \( q = \alpha/(2 - \alpha) < 1 \). In particular, for \( \alpha < 1 \), as \( h \) goes to infinity, \( q^h \to 0 \) and \( E(S) \) approaches the order of magnitude, \( l^{1-\alpha} \). This is disappointing, since the exponent is bounded from below.

To compute the average normalized waiting time when the slowdown is optimized, we first assume \( \alpha > 1 \). When we optimize slowdown all the contributions \( E_j(S) \) have bounded ratios, and running the previous argument in reverse shows that only \( E_h(N) \) will contribute. We have

\[
s_{h-1} \sim c(s_1^{\frac{q^{h-1}}{q^h-1}}) \sim d(l^{\frac{q^{h-1}}{q^h-1}})
\]

for some constants \( c, d \). Let \( \tilde{r} = \frac{q^{h-1}}{q^h-1} \), then \( \tilde{r} < 1 \). We obtain

\[
E(N) \sim E_h(N) \sim c(s_{h-1}^{2-\alpha}) \sim d(l^{-\alpha}l^{2-\alpha})
\]

For all \( \alpha > 1 \), \( \tilde{r} \) tends to \( \frac{1}{q} = \frac{2-\alpha}{\alpha+1} \), as \( h \) tends to infinity, therefore \( E(N) \) tends to \( l^{(2-\alpha)(1-\frac{\alpha}{\alpha+1})} \) as \( h \) goes to infinity. Again, this is disappointing since the exponent does not go to zero.

For \( \alpha < 1 \), we similarly have \( E(N) \sim c(s_{h-1}^{2-\alpha}) \sim d(l^{-\alpha}l^{2-\alpha}) = f(l^{(1-\tilde{r})\alpha}) \), for some constants \( c, d, f \). For \( 1/2 < \alpha < 1 \), as \( h \) tends to infinity \( 1 - \tilde{r} \) tends to \( \frac{2-\alpha}{\alpha+1} \) and so \( E(N) \) tends to \( l^{(2\alpha-1)(\frac{\alpha}{\alpha+1})} \), again a disappointment.

On the other hand, for \( \alpha < 1/2 \), \( \tilde{r} \) tends to 1 as \( h \) tends to infinity, so the average waiting time becomes an arbitrarily small power of \( l \).
Let’s summarize the main point of these tedious calculations. From these computations we conclude that for $\alpha > 1$ it is better to tune the cutoffs to provide optimal average normalized waiting time since they also lead to small slowdown, while optimizing slowdown does not lead to small waiting times. On the contrary, for $\alpha < 1/2$, it is better to tune the system for optimal slowdown for the same reasons. For the range of values $1/2 < \alpha < 1$, optimizing $E(N)$ leads to relatively poor results for $E(S)$ and vice versa. This is due to the conflicting duality properties of the objective functions in that range. We see that duality theory is at the core of the results!

8.1. Finding good cutoffs. One of the major problems in SITA queues is finding the optimal cutoffs. While the optimal cutoffs for any given instance can be computed, at least numerically, we would like a good general rule that we can follow.

Given a load $\rho$, a job size density $f$, and $h$ hosts, let $s_1^{qbal}, \ldots, s_{h-1}^{qbal}$ be the cutoffs which balance average queue length at the hosts. In some cases it may not be possible to find cutoffs which completely balance (equalize) average queue length. In those cases, we let $s_i^{qbal}$ be those cutoffs which minimize the maximal average queue length among all hosts. We have the following simple and basic result.

**Theorem 67.** For any, $s_1, \ldots, s_{h-1}$ we have

$$hE(N)(f, h, \rho, s_1, \ldots, s_{h-1}) \geq E(N)(f, h, \rho, s_1^{qbal}, \ldots, s_{h-1}^{qbal})$$

**Proof:** Given any set of cutoffs $s_1, \ldots, s_{h-1}$, let $p_i$ denote the portion of jobs which arrive at host $i$ of the SITA system, namely, the probability that $s_{i-1} \leq s < s_i$. We have $E(W) = \sum_i p_i E(W_i)$ where $E(W_i)$ is the average waiting time at the $i$’th host. For the same set of cutoffs we may also define $M(s_1, \ldots, s_{h-1}) = \max_i p_i E(W_i)$. Obviouisly, we have for any set of cutoffs $M(s_1, \ldots, s_{h-1}) \leq E(W)(s_1, \ldots, s_{h-1}) \leq hM(s_1, \ldots, s_{h-1})$. We note that the arrival rate to host $i$ is $\lambda p_i$, where $\lambda$ is the system arrival rate. Let $E(Q)_i$ be the average queue length at host $i$. By Little’s law, (81), we have that

$$E(Q)_i = \lambda p_i E(W)_i$$

By definition $s_1^{qbal}, \ldots, s_{h-1}^{qbal}$ minimizes the function $\max_i E(Q)_i$ which by (86) is the same as minimizing $M$.

We have

$$E(W)(s_1, \ldots, s_{h-1}) \geq M(s_1, \ldots, s_{h-1}) \geq M(s_1^{qbal}, \ldots, s_{h-1}^{qbal}) \geq E(W)(s_1^{qbal}, \ldots, s_{h-1}^{qbal}) / h$$

Multiplying both ends of the equation by $h/E(X)$ we get the desired result. $q.e.d.$

The result above shows that cutoffs which balance queue length are never far from optimal when minimizing average waiting time, though, the average queue lengths for the optimal cutoffs can be highly unbalanced. This occurs
for bounded Pareto distributions $P_{\alpha_n, l_n}$ with $\alpha_n \to 2$ from below and $l_n \to \infty$ at a fast enough pace. The same sequence shows that the constant $h$ is in general, the best possible. On the bright side, since we used only Little’s law the result is very general and does not require Poisson or even independent arrivals. Moreover the criterion is easy to check in practice. Even if the workload is dynamic we may try to keep the queue lengths balanced, by dynamically adjusting the cutoffs.

9. Analysis of the large $h$ asymptotics of SITA queues

As $h$ becomes large our guarantee from theorem 67 that the queue length balanced cutoffs perform well becomes very weak. We show that despite this concern, if we fix the job size distribution and increase the number of hosts, the queue length balancing heuristic for choosing cutoffs actually become near optimal when trying to minimize waiting time. It is shown in [193], [194], that the same cutoffs also minimize the probabilities of waiting a large amount of time.

We consider the case where $X_p$ is some fixed distribution of job sizes in the range $[1, p]$ with a smooth strictly positive density function $f_p$, the load $\rho$ is fixed and the number of hosts $h$ tends to infinity (becomes large). This asymptotic regime, was first considered in [85] under the name, the server expansion metric.

**Definition 131.** Let $[a, b]$, be some fixed interval. Let $g : [a, b] \to [1, p]$ be an increasing function with $g(a) = 1$ and $g(b) = p$. We will think of such functions $g$ as providing a formula for choosing the cutoffs. We say that $1 < s_{i,h} < \ldots < s_{h-1,h} < p$ is a compatible family of cutoffs, if there is a function $g$ as above such that

$$(187) \quad s_{i,h} = g(a + (b - a)i/h)$$

Let $E(N)(g, X_p, h, \rho)$ be the mean normalized waiting time for a SITA system with generic job size distribution $X_p$, load $\rho$, $h$ hosts and cutoffs given by formula (187).

**Definition 132.** We say that $g$ is an $h$-asymptotically optimal cutoff formula if for any other family of cutoffs $s_{i,h}$, compatible or not, for any $\rho$ and for any $\varepsilon > 0$ we have $E(N)(g, X_p, h, \rho)/E(N)(s_{i,h}, X_p, h, \rho) < 1 + \varepsilon$ for $h$ large enough.

The following theorem characterizes the $h$-asymptotic performance of a SITA queue.

**Theorem 68.** Consider a SITA queue, with $h$ identical hosts, a fixed utilization $\rho$ and with a job size distribution given by an infinitely differentiable density function $f$ on the interval $[1, p]$. There exists a unique, differentiable, strictly increasing function $g$ which is an $h$-asymptotically optimal cutoff formula. The function $g$ is a solution to the differential equation

$$(188) \quad f(g(x))g(x)g'(x) = c$$
where $c > 0$ is some constant, and $g$ satisfies the boundary conditions $g(a) = 1$ and $g(b) = p$. The equation says that in the $h$-asymptotic regime the asymptotically optimal cutoffs balance the queue length in each server and also balance load. The same conclusions hold for the case in which the load $\rho$ grows linearly with the number of hosts.

We have

$$E(N)^{opt} \sim \frac{\rho}{2h(1-\rho/h)}(b-a)^2c_1 \frac{1}{(m_1')^2}$$

**Proof:** Each host is responsible for job sizes in a small range given by equation (187). Let $dx_i = (b-a)/h$. Let $ds_i = s_i - s_{i-1}$. We have

$$ds_i = s_i - s_{i-1} = g(a + (b-a)i/h) - g(a + (b-a)(i-1)/h) \sim g'(x_i)(b-a)/h = g'(x_i)dx_i$$

Since the density function is bounded from above, we see that the load $\rho_i$ on each host tends to zero and hence $1 - \rho_i \to 1$. We conclude that

$$E(N_i) \sim \frac{\rho L_{f,I_1}(1)L_{f,I_1}(3)}{(L_f(2))^2}$$

We have

$$L_{f,I_1}(3) \sim g'(x_i)f(g(x_i))g(x_i)^2dx_i$$

Similarly

$$L_{f,I_1}(1) \sim g'(x_i)f(g(x_i))dx_i$$

Summing over $i$ and recalling that $dx_i = \frac{b-a}{h}$ we have the Riemann sum expression

$$E(N) = \sum_i E(N_i)$$

$$\sim \frac{\rho}{2} \sum_i (f(g(x_i))g(x_i))^2g'(x_i)^2dx_i^2$$

$$= \frac{\rho(b-a)}{2h(L_f(2))^2} \sum_i (f(g(x_i))g(x_i))^2g'(x_i)^2dx_i$$

Taking the limit as $h$ goes to infinity we see that

$$E(N) \sim \frac{\rho(b-a)}{2h(L_f(2))^2} \int_a^b (f(g(x))g(x))^2g'(x)^2dx$$

We need to minimize the integral expression. Given $h$ hosts and optimal cutoffs $s_1, \ldots, s_{h-1}$ we define as described before, the non decreasing curve $g_h$ by extrapolating linearly between the cutoff values. Since the set of non decreasing curves is compact in the $C_0$ topology, a subsequence of the curves $g_h$ converges in that topology to a curve $g_\infty$.

We note that the functional of interest has the form $F(x, g, g') = \tilde{L}(g)g'^2$, which is a special case of the more general form more

$$F(x, g, g') = \tilde{L}(x, g)g'^2$$
This bares a strong resemblance to functionals of the form
\[ F(x,g,g') = \tilde{L}(x,g)\sqrt{g'} \]
which we considered in chapter 3 as the geodesic equation in a two dimensional space-time with a Lorentzian metric. The main difference between these functionals is that the function \(z^2\) which is applied to \(g'\) in the first case is strictly convex, while the function \(\sqrt{z}\) which is applied to \(g'\) in the second case is strictly concave. This difference in particular is responsible for the fact that a generic solution to the Euler-Lagrange equation in the first case is a minimum while in the second case it is a maximum. Following the same arguments as in chapter 3 the functional is lower semi-continuous in the \(C_0\) topology, hence \(g_\infty\) is a minimizer and consequently an \(h\)-asymptotically optimal cutoff formula exists. The arguments that the limiting curve \(g_\infty\) is the graph of a differentiable strictly increasing function come again from the theory of existence and uniqueness of ODE, see [61] or [142] for similar arguments. Since \(x\) does not figure explicitly in \(F = \tilde{L}(g)g'^2 = f^2(g)g^2g'^2\) the Euler-Lagrange equation which a minimizing function must satisfy degenerates to the Beltrami equation as presented in theorem 28
\[ g\frac{\partial L}{\partial g'} - L = c_1 \]
where \(c_1\) is some constant, or in this particular case
\[ 2g'^2\tilde{L} - \tilde{L}g'^2 = \tilde{L}g'^2 = F(g,g') = c_1 \]
The constant \(c_1\) must be positive because \(L\) and \(g'\) are. The equation (188) is obtained by taking the square root. Since as \(h\) becomes large \(E_i(N) \sim F(g,g')(a + i(b - a)/h)\) we see that we may interpret the Beltrami equation as load balancing these contributions, which by Little’s law also means balancing the average queue lengths at the hosts. The equation (188) itself describes the relative load at each host, hence this is also load balanced.

Plugging the equation into the computation of \(E(N)\) and recalling that \(\rho/h \to 0\) when \(\rho\) is fixed, gives the expression in (189). The solution is unique since there is a unique load balancing solution for each \(h\). We can also argue that it is unique since it also solves the geodesic equation in Minkowski space.

If we consider the case where the load \(\rho = \tilde{\rho}h\) grows linearly with \(h\) we obtain a slightly more complicated functional
\[ \int_a^b \frac{f(g(x))g(x)^2g'(x)^2}{1 - \rho(f(g(x))g(x))} dx \]
but, the solution to the corresponding Beltrami equation is the same as before. \(q.e.d.\)

We may consider linearly coupled servers in the context of theorem 68. We will need to assume that the speed up constant \(s_{i,h}\), comes from a smooth
function \( s(x) \) by placing \( s_i = s(a + i(b - a))/h \). This amounts to a restriction that hosts which handle nearby job size intervals have nearly equal processing speed. Under this assumption the corresponding functional for determining asymptotically optimal cutoffs will have the general form

\[
A(x)B(g)(g')^2
\]

Since we have a product form \( A(x)B(g) \), we can make separate monotone changes of variables to \( x \) and \( g \) (or \( y \) in the usual coordinate notation), which reduces the functional to the identical server case. We get after the change of coordinates the previous results about uniqueness of the solution and having a conserved quantity (A Beltrami equation). We note that the same holds in Lorentzian geometry, the equation \( A(x)B(g)\sqrt{g'} \) is the geodesic equation of Minkowski space, but in non-standard coordinates where one applies monotone changes of variables to the \( x \) and \( y \) coordinates separately. Essentially we are still in the Minkowski space case.

We may consider, more generally, the case of (non-linearly) coupled servers. To consider the large \( h \) limit, we again assume that the functions \( T_{i,h}(t) \), lead in the limit to a continuous function \( T(x,g) \). The resulting functional will have the form

\[
\tilde{L}(x,g)(g')^2
\]

This is similar to the proper time functional in Lorentzian geometry, however, this time the Lorentzian metric is not Minkowski, it is in general curved space-time, which no coordinate change will convert to Minkowski space. In analogy with Lorentzian geometry, we do not expect in this case to always have a unique solution and examples can be constructed where the solution is not unique. The solution need not be unique since the Euler-Lagrange equation is not a Beltrami type equation anymore and we do not have a balanced (conserved) quantity. We conclude that going from the linearly coupled model to the coupled model is analogous in this case to the introduction of curvature in space-time or to passing from special relativity to general relativity.

Notes for chapter 5

(1) We note that the discussion in sub-section 3.2 explains one of the great mysteries in life, namely, why the bus never shows up on time. Assume that we arrive at some random point in time to the initial bus station in the bus route. Assume that the bus departs every 15 minutes. Since we arrived at some random time point we expect to wait on average 7.5 minutes and this is consistent with equation (108) when we enter the deterministic distribution with value 15. On the other hand if we arrive to a station somewhere in the middle of the route and there is heavy traffic along the way, the traffic will make the arrivals at the station more erratic, but obviously (unless buses get lost) still with an average of 15 minutes. Assume that
traffic makes the arrivals Poisson with an average of 15 minutes, then plugging into the equation we get that we have to wait 15 minutes on average! This makes sense because a Poisson process is memoryless so you wait 15 minutes regardless of how long ago the last bus arrived. Things can obviously get even worse if the inter-arrival distribution has even larger variance.

(2) Our first example of a self-dual function of number theoretic origin was the theta function. We also saw that Riemann showed that the Mellin transform of the theta function is closely related to the zeta function. We explain how this can be extended via an operation which is known as twisting by a character. This construction will generate more examples of self-dual functions from the initial theta function.

**Definition 133.** Given a primitive character $\chi$ of conductor $N$ (see appendix), we may define the Gauss sum

$$\tau_\chi = \sum_{Z_N} \chi(n)e^{2\pi in/N}$$

It can be shown that $|\tau_\chi|^2 = N \neq 0$ and that, see [45],

$$\chi(n) = \frac{\chi(-1)\tau_\chi}{N} \sum_{m \in (Z_N)^*} \bar{\chi}(m)e^{2\pi imn/N}$$

During his study of primes in arithmetic progressions, [65], Dirichlet considered a version of the zeta function, twisted by a character.

**Definition 134.** Let $\chi$ be a character with conductor $N$. The Dirichlet $L$-function associated with $\chi$ is given by

$$L_\chi(s) = \sum_{N=1}^{\infty} \chi(n)n^{-s}$$

The Riemann zeta function corresponds to the trivial character with $N = 1$. Since a character is a multiplicative homomorphism we have the Euler product form

$$L_\chi(s) = \prod_p (1 - \frac{\chi(p)}{p^s})^{-1}$$

when $\text{Re}(s) > 1$.

**Definition 135.** Assuming $\chi$ is primitive, $\chi(-1) = 1$ and that $N > 1$ we define the twisted $\theta$ function

$$\theta_\chi(t) = \frac{1}{2} \sum_{n=-\infty}^{\infty} \chi(n)e^{-\pi n^2t} = \sum_{n=1}^{\infty} \chi(n)e^{-\pi n^2t}$$
If $\chi(-1) = -1$ (odd character), then the terms for $n$ and $-n$ above will cancel. Instead, one uses the definition

$$\theta_\chi(t) = \sum_{n=1}^{\infty} n\chi(n)e^{-\pi n^2 t}$$

(194)

To treat the cases together we define $\varepsilon = \varepsilon_\chi = 0, 1$ by the formula $\chi(-1) = (-1)^\varepsilon$.

Repeating the arguments leading to the Poisson summation formula in conjunction with equation (190) leads to the twisted functional equation

$$\theta_\chi(t) = -i\tau(\chi)N^{-1-\varepsilon}t^{-1/2-\varepsilon}\theta_\chi(\frac{1}{N^2 t})$$

(195)

**Definition 136.** We define the completed twisted L-function of a character by the formula

$$\Lambda_\chi(s) = \pi^{-(s+\varepsilon)/2}\Gamma\left(\frac{s+\varepsilon}{2}\right)L_\chi(s)$$

(196)

In analogy with the zeta function calculation it can be shown that

$$L_{\theta_\chi}((s+\varepsilon)/2) = \Lambda_\chi(s)$$

(197)

The exponential decay of $\theta_\chi$ as $t \to \infty$ and the functional equation (195) which proves exponential decay at 0 show that $\Lambda_\chi$ extends to a holomorphic function with no poles on the entire complex plane. In addition, we obtain the functional equation

$$\Lambda_\chi(s) = (-i)^\varepsilon \tau(\chi)N^{-s}\Lambda_\chi(1-s)$$

(198)

The same type of twisting can be performed on modular forms.

**Definition 137.** Let $f \in S_k^1(N, \psi)$ for some Dirichlet character $\psi$ and let $f = \sum_n f_n q^n$ be its Fourier expansion. Let $\chi$ be a primitive Dirichlet character modulo $D$, with g.c.d.($D, N$) = 1. We define the twisting of $f$ by the character $\chi$ to be the function $f^\chi$ with expansion

$$f^\chi = \sum_n \chi(n)f_n q^n$$

Using (190) we see that

$$\chi(n)q^n = \chi(n)e^{2\pi i nz}$$

$$= \frac{\chi(-1)}{D} \sum_{m \in (Z/DZ)^*} \bar{\chi}(m)e^{2\pi in(z+m/D)}$$

We conclude that $f^\chi$ is a combination of $f_\delta(D, m)$ where $\delta(D, m)$ is the element

$$\left( \begin{array}{cc} D & m \\ 0 & D \end{array} \right)$$

(199)
for which $\delta(D, m)(z) = z + m/D$. It is easy to verify that

$$\Gamma_1(D^2N) \subseteq \delta(D, m)^{-1}\Gamma_1(N)\delta(D, m)$$

for all $m$ and so $f^\chi \in S_k^1(D^2N)$ and in fact $f^\chi \in S_k^1(ND^2, \chi^2\psi)$.

Furthermore using the above it can be shown that restricting to the positive imaginary line

$$(200) \quad H_{k, ND^2}(f^\chi) = \chi(N)\psi(D)^{\frac{\tau_2}{D}}H_{k, N}(f^\chi)$$

From the above discussion, we conclude that all the twists of a form in $S_k^1(N, \psi)$ are related via the involutions $H_{k, M}$ for various $M$ to all the twists of the dual form. We may ask if in analogy with the case of log-normal distributions having these many functional duality relations characterizes such forms. We can now state Weil’s converse theorem, [184], which answers this question positively.

**Theorem 69.** Let $N$ be a positive integer and $\psi$ a Dirichlet character of $(\mathbb{Z}/N)^*$. Let $f = \sum f_n q^n$ and $g = \sum g_n q^n$, where $f_n, g_n$ are some sequences of polynomial growth, namely, $|f_n|, |g_n| \leq cn^r$, for some $c, r > 0$. If for any primitive Dirichlet character $\chi$ of conductor $D$, which satisfies $\text{g.c.d.}(D, N) = 1$, the functions $L_{f, \chi}$ and $L_{g, \bar{\chi}}$ can be analytically continued to the entire plane, are bounded in every vertical strip $A \leq \text{Re}(s) \leq B$ and satisfy the functional equation (200), then, $f \in M_k^1(N, \psi)$.

Weil’s converse theorem shows the tight relation between the involutions $H_{k, N}$ which are of queueing theoretic significance and modularity of Fourier (Taylor) expansions in the variable $q$. The only issue is that the functions $f^\chi$ will (in general) not be positive or even real. It would be interesting if we could find a ”real” version of Weil’s theorem. For example, what can we recover if we know equation (200) at the level of absolute values on the positive imaginary axis. Since $|\chi(N)\psi(D)^{\frac{\tau_2}{D}}| = 1$, we would lose the information about $f$ being in $M_k^1(\psi)$. Could we still show that $f$ is modular, knowing also that all the functions $f^\chi$ are analytic.

(3) We see that the twisted $\theta$ functions have a significant number theoretic role, if they happen to be positive, then by the functional equation, they will have nice queueing theoretic properties. There is also number theoretic interest in positive $g$. A generalization of the Riemann conjecture asserts that all the zeroes of $L_{\theta_p}$ will lie on the line $\text{Re}(s) = 1/2$ which is fixed under the symmetry of the functional equation. If $\theta_p$ is positive then we know at least that $L_{\theta_p}$ will not have zeroes on the real line, this has already been observed by Riemann himself. In addition, they have interesting moments. The combination of these properties makes them interesting as density
functions. To be positive, we want the values of $\chi$ to be real. Since these values are roots of unity, we want the values to be 1, −1.

**Definition 138.** Characters such that $\chi^2 = 1$ are called **quadratic characters**.

Each prime $p > 2$ has a quadratic character $\chi_p$, which is simply the quadratic residue modulo $p$, i.e., $\chi_p(n) = 1$ if and only if $g.c.d.(p, n) = 1$ and the equation $x^2 = n \pmod{p}$ has an integer solution, see chapter 4 for more on this. This turns out to be the only non-trivial quadratic character modulo $p$, since the multiplicative group modulo $p$ is cyclic and so the character is determined by the value on a generator which must be −1, if the character is to be non-trivial. We can ask whether $\theta_p = \theta_{\chi_p}$ is a positive function. This question is very old and has been studied extensively. We first consider a reformulation. If the function $\theta_p$ is positive then so is its Laplace transform

$$F_p(\lambda) = \int_0^\infty e^{-\lambda t} \theta_p(t) dt$$

Using the definitions (193,194) and the functional equation (195) we get the following calculation of $F_p$

$$F_p = \sum_{n=-\infty}^{\infty} \int_0^\infty \frac{n\chi(n)}{(Nt)^{3/2}} e^{-\lambda t} e^{-\pi n^2/(N^2 t)} dt$$

$$= \frac{1}{\sqrt{N}} \sum_{n=1}^{\infty} \chi(n) e^{-(2n\sqrt{\pi}\lambda/N)}$$

We write $x = e^{-2n\sqrt{\pi}\lambda/N}$. As $\lambda$ varies over the positive reals, $x$ varies in the range $0 < x < 1$. Since $\chi$ is periodic with period $N$ we get

$$F_p = \frac{1}{\sqrt{N(1-x^N)}} \sum_{n=1}^{N-1} \chi(n)x^n$$

From this formula we see that the Laplace transform is positive if and only if for all $0 < x < 1$, the polynomial

$$f_\chi(x) = \sum_{n=1}^{N-1} \chi(n)x^n$$

is positive. These polynomials are known as the Fekete polynomials, even though they were already introduced by Dirichlet.

Apparently, see [145], Fekete once conjectured that all Fekete polynomials are positive in the range $0 < x < 1$. This was shown to be false by G. Polya who proved that for $p = 43$ we have $f_{\chi_p}(0.7) < 0$. However, for small values of $p$ most $f_p$ are positive, in fact,
according to \[24\] only 23 primes less than 1000 have non-positive $f_p$.

Polya’s argument runs as follows, assume that $\chi_p$ is such that
\[
\chi_p(2) = \chi_p(3) = \chi_p(5) = \chi_p(7) = \chi_p(11) = -1
\]
Using multiplicativity we can compute the values of $\chi_p(n)$ for $n \leq 12$. since $\chi_p(n) \leq 1$ we have in that case
\[
f_p(x) \leq x - x^2 - x^3 + x^4 - x^5 + x^6 - x^7 - x^8 + x^9 + x^{10} + x^{11} - x^{12} + \frac{x^{13}}{1-x}
\]
The left hand side is negative in the interval [0.69, 0.8]. Note that using quadratic reciprocity and the Chinese remainder theorem, we can translate the condition in (201) to a condition that $p$ modulo 8·3·5·7·11 = 9240 is in a certain subset $A \subset (\mathbb{Z}_{9240})^*$ which includes 1/32 of all residues (five independent conditions, each satisfied by half the residues). By the quantitative version of Dirichlet’s theorem, in the limit, a proportion of 1/32 of all primes will satisfy all five conditions and hence will have non positive $\theta_p$. More conditions that guarantee that $\theta_p$ will not be positive were found in \[83\]. Finally, it was proved in \[22\] that for any $k > 0$, for almost all primes $f_p$ has more than $k$ zeroes in the range $0 < x < 1$, hence having positive $\theta_p$ is unfortunately an asymptotically rare phenomenon. It was asked in \[83\] and again in \[22\] whether the number of positive $f_p$ is finite. This remains an open problem. For the same reasons that we inquired about positivity of the $\theta_p$ we can ask whether Hecke newforms, and in particular the functions $g_{A,B}$ are positive. The moments of these functions are conjecturally of great arithmetic interest and hence also the moments of waiting time in an M/G/1 queue with such densities, see \[26, 27, 31, 32, 100\].

As an example of the arithmetic interest in moments, consider the basic step of converting $g = g_E = g_{A,B}$ from a generalized density into an actual probability density. The function $g_E$ was normalized in such a way that $a_1 = 1$. This makes sense from a number theory point of view since it made the coefficient $a_n$ coincide with the eigenvalue of the Hecke operator $T_n$. However, for the purposes of probability theory we need to divide $g_E$ by the constant $\int_0^{\infty} g_E(t) dt$. By definition, this constant is $L_\varrho(1) = L_E(1)$. The arithmetic interpretation of this normalizing term is the subject of the Birch and Swinnerton-Dyer conjecture (BSD), \[30\]. This conjecture relates the arithmetic properties of the set of solutions of the equation $y^2 = x^3 + Ax + B$ over various fields to the value $L_E(1)$. Note that this is the mid point on the real line for the functional equation which relates $L_\varrho(s)$ with $L_\varrho(2 - s)$.
(recall that \( k = 2 \) for \( g = g_E \)). To discuss the value \( L_E(1) \), we need to add to the solutions of \( p(x,y) = 0 \) solutions at infinity, in analogy with the cusps of a fundamental domain. This is a standard procedure in algebraic geometry in which we add one variable \( z \) to \( x \) and \( y \) and multiply all the monomials of \( p(x,y) \) by the appropriate power of \( z \) so that in the resulting polynomial \( \tilde{p}(x,y,z) \), all monomials have the same degree. Applied to \( p_{A,B} \) we get \( \tilde{p}_{A,B} = zy^2 - x^3 - Axz^2 - Bz^3 \). We then consider the solutions of \( \tilde{p}(x,y,z) = 0 \) in projective space, namely we consider solutions \((x,y,z) \neq (0,0,0)\), under the equivalence relation that identifies \((x,y,z)\) with \((cx, cy, cz)\) for any \( c \). If \( z \neq 0 \) in a solution then by multiplying the solution by \( c = z^{-1} \) we get an equivalent solution with \( z = 1 \) and plugging into \( \tilde{p}(x,y,z) = 0 \), we see that we get a solution to the original equation \( p(x,y) = 0 \). The added solution at "infinity" corresponds to the case \( z = 0 \), which after plugging into the equation yields the condition \( x = 0 \). Multiplying by a constant we get the unique (up to equivalence) additional solution \((0,1,0)\).

We still denote the extended set of solutions by \( E \) and call it an elliptic curve, in fact, this is the correct definition (in a very naive form). We note that any rational function, i.e., a function of the form \( F(x,y,z) = r(x,y,z)/s(x,y,z) \), where \( r, s \) are polynomials all of whose monomials have the same degree respects the equivalence relation and is thus well defined on \( E \). It is known that there are rational functions with coefficients in \( \mathbb{Q} \) that define a commutative group structure on \( E \). See [163] for the explicit construction. By a fundamental theorem of Mordell, [136], the set of solutions with coordinates in \( \mathbb{Q} \), denoted \( E(\mathbb{Q}) \), is a finitely generated Abelian group, see [163] for an elementary proof. By the theory of finitely generated Abelian groups the group \( E(\mathbb{Q}) \) has the form \( T \oplus \mathbb{Z}^r \), where \( T \) is a finite group consisting of the finite order elements (torsion elements). The number \( r \) is a well defined invariant of the group.

**Definition 139.** If \( E \) is an elliptic curve with integer coefficients then the number \( r = r_E = r_{E}^{alg} \) is known as the (algebraic) rank of the elliptic curve.

If we consider the complex analytic function \( L_g \), it has a certain order of vanishing at 1, namely, we may write \( L_g(1) = (s - 1)^{r_{anal}} \tilde{L}_g(s) \), where \( \tilde{L}_g \) is non vanishing at \( s = 1 \).

**Definition 140.** The integer \( r_{anal} \) is known as the analytic rank of \( E \).

A weak form of the BSD conjecture makes the fantastic claim that \( r_{anal} = r_{alg} \). In [117], Kolyvagin proved the conjecture for the cases where \( r_{anal} \) is either 0 \( (L_g(1) \neq 0) \) or 1, assuming the
Taniyama-Shimura conjecture which was not known at the time. Following [41] this is true for general elliptic curves with integer coefficients.

Obviously, if $g$ is strictly positive then $L_g(1)$ will be positive, i.e., $r_{E}^{anal} = 0$. By Kolyvagin, we can hope for a positive $g_E$ only if the elliptic curve which we started with has no rational point of infinite order. This is a necessary condition, but is not sufficient.

We can test numerically whether a given curve $E$ leads to a positive function $g_E$. We first note that by the functional equation it is enough to verify positivity on the interval $[\frac{1}{\sqrt{N}}, \infty]$. Moreover, we have the estimate $|a_n| \leq n$. Taking a factor of $q$ out of $g_E$ we see that $g_E$ is positive if and only if $1 + \sum_{k=2}^{\infty} a_k q^{k-1}$ is positive a condition which is guaranteed if $\sum_{k=2}^{\infty} k q^{k-1} = \frac{1}{(1-q)^2} - 1 \leq 1$, which leads to $q \leq 1 - \sqrt{\frac{1}{2}}$ or $t \geq -\log(1-\sqrt{\frac{1}{2}}) = 0.195\ldots$. Consequently, it is enough to check positivity of $g_E/q$ on the interval $[\frac{1}{\sqrt{N}}, 0.196]$. This means that all curves with $N < 27$ automatically have positive $g_E$ functions. The first case of an elliptic curve which is not positive occurs when $N = 37$ for the simple reason that it has rank 1. The first case of a non positive $g_E$ when $r_E = 0$ occurs when $N = 203$.

The mapping from elliptic curves to functions $g_E$ is not 1-1. A deep result of Faltings, [71, 54] says that two elliptic curves $E_1$ and $E_2$ will produce equal functions $f_{E_1} = f_{E_2}$ if there is a mapping $E_1 \mapsto E_2$ (or from $E_2$ to $E_1$) given by rational functions with coefficients in $\mathbb{Q}$, such that the inverse image of a point is finite. Such curves are called isogeneous.

Up to isogeny, there are 1321 rank zero elliptic curves over the rational numbers, with conductor $N < 1000$. Together with A. Bond and S. Ryak, we checked positivity for all these functions and found that 1219 out of 1321 were positive. In the range $100,000 < N_E < 100,300$, we found 336 out of 441 rank 0 cases to be positive, which is obviously a lower percentage but still a majority.

However, as in the case of the twisted theta functions it has been proved very recently, see [103] only a sparse set of values $A, B$ have positive $g_{A,B}$. In fact this is proved for many families of modular forms, including the Hecke newforms of a given weight with varying levels $N$. When we fix the level $N$ and let $k \to \infty$, it has recently been shown, [81], that only a finite number of Hecke newforms $g$ will be positive on the imaginary axis. See also [80] for a related result. For a more comprehensive account of these developments and the state of the art regarding the question whether there is an infinite number of Hecke newforms which are positive, see [156, 157].
Appendix A: Some basic definitions and facts

1. Point-set topology and metric topology

Definition 141. Let $X$ be a set. A family $T$ of subsets of $X$ is called a topology on $X$ if

(i) The empty set and $X$ itself belong to $T$.

(ii) The intersection of finitely many elements in $T$ is also in $T$.

(iii) Any union of elements in $T$ is in $T$.

The elements in $T$ are called the open sets of the topology and their complements are called the closed sets. The pair $(X, T)$ is called a topological space.

A family $B \subseteq T$ of open sets is called a basis for the topology $T$ if, for any point $x \in X$ and for any open set $U \in T$ such that $x \in U$, there is a $V \in B$ such that $x \in V \subset U$.

Definition 142. If $(X, T)$ is a topological space and $A \subseteq X$, the closure of $A$, denoted $\overline{A}$, is the smallest closed set containing $A$, or equivalently, the intersection of all closed sets containing $A$. The boundary of $A$ is given by $\overline{A} \cap \overline{X} - A$.

If $(X, d)$ is a metric space, then the metric $d$ induces a topology $T_d$ on $X$. A set $U$ is open in $T_d$ if and only if, for any $x \in U$ there exists an $\varepsilon_{x, U}$ such that $B(x, \varepsilon_{x, U}) \subseteq U$. Here, $B(x, \varepsilon)$ denotes the open ball of radius $\varepsilon$ with center $x$, i.e., the set of points $y$ such that $d(x, y) < \varepsilon$. It is easy to check that this defines a topology and that the set of all open balls forms a basis for $T_d$.

We note that different metrics may define the same topology. For example, all the $\ell_p$ metrics on $\mathbb{R}^n$ define the same topology, as can be easily verified.

Definition 143. We say that a topological space $(X, T)$ is a (topological) metric space if there exists a metric $d$ on $X$ such that $T = T_d$.

We collect a few more relevant definitions.

Definition 144. Let $(X, T_X), (Y, T_Y)$ be topological spaces. A function $f : X \to Y$ is said to be continuous if for every $U \in T_Y$ we have $f^{-1}(U) \in T_X$. 

201
The spaces are said to be **homeomorphic** if there exist continuous functions \( f : X \to Y \) and \( g : Y \to X \) which are inverse to each other.

**Definition 145.** If \((X,T_X)\) is a topological space and \(Y \subset X\) then we get an **induced topology** on \(Y\) by letting \(W \subset Y\) be an open set if and only if there exists \(U \in T_X\) such that \(W = U \cap Y\).

**Definition 146.** A topological space is **compact** if for every family of open sets \(U_\alpha, \alpha \in A\) such that \(\bigcup \alpha \in A U_\alpha = X\) there exists a finite \(B \subset A\) such that \(\bigcup \alpha \in B U_\alpha = X\). In other words, every open cover of \(X\) has a finite sub-cover.

**Definition 147.** Let \((X,T_X)\) be a topological space and let \(x_i \in X, i = 1, 2, ...\) be a sequence of elements. We say that the sequence \(x_i\) **converges** to an element \(x \in X\), if for every open set \(U\) such that \(x \in U\), there exists an \(N_U\) such that for all \(n > N_U\), \(x_n \in U\). We say that \(x\) is a limit point of a sequence \(x_i\), if there is a subsequence \(x_{i_j}\) of \(x_i\) which converges to \(x\).

In this generality, a sequence may converge to more than one point. Indeed, if \(X\) is any set and \(T_X\) is the trivial topology which contains only the empty set and \(X\) then any sequence converges to any point. However, it is easy to check that, in metric space convergence, if it exists is to a unique point.

**Definition 148.** We say that a topological space is **sequentially compact** if any sequence \(x_i\) has a convergent sub-sequence \(x_{i_j}\).

In general topological spaces compactness does not imply sequential compactness and vice versa. However, in metric spaces the two notions coincide. To see this, we begin with the following lemma

**Lemma 19.** Let \((X,T_X)\) be a sequentially compact space, then any continuous function \(f : X \to \mathbb{R}\) has both a minimum and a maximum.

**Proof:** Let \(m\) be the infimum of \(f\) on \(X\), which in principle could also have the value \(-\infty\). By definition, we can find an \(x_i\) such that \(f(x_i) < m + 1/i\). Since \(X\) is sequentially compact, the sequence \(x_i\) has a convergent subsequence to some element \(x \in X\). From the definition of continuity, we must have \(f(x) = m\) since otherwise, if \(f(x) = \tilde{m} > m\) then the open set \(f^{-1}(B(\tilde{m}, (\tilde{m} - m)/2) \text{ or } B(\tilde{m}, 1) \text{ if } m = -\infty)\) contains only a finite number of \(x_i\) and hence no subsequence can converge to \(x\). We conclude that \(m\) is finite and is attained. The proof for the maximum is the same (or apply the argument to \(-f\)). \(q.e.d.\)

**Lemma 20.** Let \(X\) be a metric space, then the closure of a sequence of points \(x_i\) consists of the sequence \(x_i\) and its limit points.

**Proof:** If \(x\) is not in the closure of \(x_i\), then it is contained in an open set which is disjoint from the sequence, hence it is not a limit point. If it is in the
closure, but is not an element of the sequence then any open set containing
\(x\) will contain a point of the sequence. We can then inductively produce a
convergent subsequence by choosing the \(j\)'th element of the subsequence in
the open ball of radius \(min(1/j, d(x, x_1, ..., d(x, x_{j-1}))\). q.e.d.

**Theorem 70.** For topological metric spaces the notions of sequential
compactness and compactness are equivalent.

**Proof:** Let us first assume that a metric space \(X\) is compact. Let \(x_i \in X\)
be some sequence of distinct points with no convergent subsequence. Let
\(A_i = \{x_j \mid j \geq i\}\). By lemma 20, the \(A_i\) are closed, hence their complements
\(B_i = X - A_i\) are open. Since the intersection of the \(A_i\) is empty, the union
of the \(B_i\) is \(X\). However, by definition there is no finite sub-cover.

In the other direction, assume \(X\) is sequentially compact. Let \(U_\alpha\) form
an open covering of \(X\). For a given \(x \in X\), let \(f(x)\) be the supremum over all
\(r\) such that there is some open set \(U_\alpha(x)\) in the cover with \(B(x, r) \subset U_\alpha(x)\).
Since \(x \in \bigcup_\alpha U_{\alpha} = X\), we have \(f(x) > 0\). Also, since the distance from
\(x\) is a continuous function on \(X\) (easy to check), it is bounded by lemma
19, so \(f(x)\) is well defined. It is easy to see that \(|f(x) - f(y)| \leq d(x, y)\),
hence \(f\) is continuous. By lemma 19 we conclude that \(f\) has a positive
minimum, say \(\varepsilon\). We consider the open cover of \(X\) consisting of the open
balls \(B(x, \varepsilon/2)\) for all \(x \in X\). We claim that this open cover has a finite
sub-cover. If not, we choose \(x_1\) arbitrarily and define inductively \(x_{i+1}\)
to be a point which is not in \(\bigcup_{j \leq i} B(x_j, \varepsilon/2)\). The distance between any two
points in the resulting sequence satisfies \(d(x_i, x_j) \geq \varepsilon/2\), hence, given some
point \(x\), the open ball \(B(x, \varepsilon/4)\) will contain at most a single point from
the sequence \(x_i\). Consequently, no \(x\) can be a limit point for the sequence
\(x_i\) and we obtain a contradiction. Given the finite sub-cover \(B(y_k, \varepsilon/2),
\(k = 1, ..., N\) and given the definition of \(\varepsilon\) there are open sets \(U_{\alpha(k)}\) such that
\(B(y_k, \varepsilon/2) \subset U_{\alpha(k)}\), hence they form the required open sub-cover. q.e.d.

**Definition 149.** A function \(\ell\) is upper semi-continuous on a topologi-
cal space \(X\) if for any \(\varepsilon > 0\) and for any point \(x \in X\), there is an open
neighborhood \(x \in N(x)\) such that for any \(x' \in N(x)\), \(\ell(x') \leq \ell(x) + \varepsilon\).

**Lemma 21.** An upper semi-continuous function \(\ell\) on a compact and se-
quentially compact space \(X\) attains a maximal value.

**Proof:** Let \(N_x\) be the open cover which is guaranteed to exist by the def-
nition of upper semi-continuity when we take \(\varepsilon = 1\). Since \(X\) is compact,
it is covered by a finite subset \(N_{x_1}, ..., N_{x_m}\). By definition, the function \(\ell\)
will be bounded from above by \(Max(\ell(x_1), ..., \ell(x_m)) + 1\). Let \(M\) be the
supremum of \(\ell\), we know that it is finite. Let \(x_n \in X\) be a sequence of
points with \(\ell(x_n) \rightarrow M\). Since \(X\) is sequentially compact, the sequence \(x_n\)
has a subsequence which converges to some value \(x\). Assume \(\ell(x) < M\).
By upper semi-continuity of \(\ell\) there is a neighborhood \(V_x\) of \(x\) such that
for \( x' \in V(x) \) we have \( l(x') < (l(x) + M)/2 < M \), and this contradicts our previous assumptions. \( q.e.d. \)

**Definition 150.** Let \( X \) be a topological space. A subset \( A \subset X \) is called dense if \( \overline{A} = X \).

**Lemma 22.** A sequentially compact space has a countable dense set.

**Proof:** As shown in the proof of theorem 1 there are finite open covers of the form \( B(x_{i,j},1/j) \), \( i = 1, ..., n(j) \) for every \( j \). Taking the union of points \( x_{i,j} \) and noting that the sets \( B(x,1/j) \) form a basis we see that the union is dense. \( q.e.d. \)

**Definition 151.** Let \( X, Y \) be metric spaces with metrics \( d_X, d_Y \) respectively. We say that a function \( f : X \rightarrow Y \) is Lifschitz with Lifschitz constant \( c \), if for any \( x_1, x_2 \in X \) we have \( \frac{d_Y(f(x_1),f(x_2))}{d_X(x_1,x_2)} \leq c. \)

There is a general construction extending the distance between points in a metric space to a distance between any two closed sets \( X, Y \) in the metric space. Given a point \( x \in X \) we define the distance between the point \( x \) and the set \( Y \) as \( d_S(x,Y) = \text{Inf}_{y \in Y} d_S(x,y) \). If \( Y \) is compact then the function \( y \rightarrow d_S(x,y) \) is continuous and the infimum is attained. We define \( \tilde{d}_S(X,Y) \) as \( \text{Sup}_{x \in X} d_S(x,Y) \). If \( X \) is compact the supremum is attained. The function \( \tilde{d}_S(X,Y) \) is usually not symmetric, thus we define

**Definition 152.** Let \( (S,d_S) \) be a metric space and \( X, Y \subset S \) closed sets. The **Hausdorff distance** between \( X \) and \( Y \) is defined by

\[
d_S(X,Y) = \text{Max}(\tilde{d}_S(X,Y), \tilde{d}_S(Y,X))
\]

A simple argument shows that \( d_S \) inherits the triangle inequality property from \( d_S \) on points of \( S \). The metric \( d_S \) restricted to pairs of causal curves defines the Hausdorff metric on the space of causal sets.

**Definition 153.** We say that a set of functions \( F \), from one metric space \( (X,d_X) \) to another \( (Y,d_Y) \) is equi-continuous if for each \( \varepsilon > 0 \) there exists a \( \delta > 0 \) such that for any \( f \in F \), if \( d_X(x_1,x_2) < \delta \) then \( d_Y(f(x_1),f(x_2)) < \varepsilon. \)

The notion of equi-continuity extends the definition of uniform continuity from a single function to a family of functions. As an example, the set of Lifschitz functions from \( X \) to \( Y \) with a given Lifschitz constant \( c \) is equi-continuous. This is seen by setting \( \delta = \varepsilon/c. \)

**Definition 154.** We say that a sequence of functions \( f_n : X \rightarrow Y \) **converges uniformly** to a function \( f \), if for each \( \varepsilon > 0 \), there exists an \( N_\varepsilon \) such that for all \( x \in X \) and all \( n > N_\varepsilon \), \( d_Y(f_n(x), f(x)) < \varepsilon. \)
Definition 155. Let $X$ and $Y$ be compact metric spaces. We define a metric on the set of all continuous functions from $X$ to $Y$ by the formula $d_{X,Y}(f,g) = \max_{x \in X} d_Y(f(x), g(x))$.

It is easy to verify that if $f$ and $g$ are continuous then the function $x \to d_Y(f(x), g(x))$ is a continuous real valued function on $X$ and hence attains a maximum. Using this fact and the triangle inequality for $d_Y$ it is easy to verify the triangle inequality for $d_{X,Y}$ and hence that it is indeed a metric. It is easy to verify from the definition that a sequence $f_n$ converges uniformly to $g$ if and only if it converges w.r.t. the metric $d_{X,Y}$.

The following is a variant of a classical result known as the Arzela-Ascoli theorem which relates equi-continuity to sequential compactness.

Theorem 71. Let $F$ be a closed family of equi-continuous functions from a compact metric space $X$ to a compact metric space $Y$, equipped with the metric $d_{X,Y}$, then $F$ is sequentially compact and hence compact. In particular, the family of Lifschitz functions with a bounded Lifschitz constant is compact.

Proof: Since $X$ is compact metric it is also sequentially compact, hence it has a countable dense set. Let $x_1, x_2, \ldots, x_n, \ldots$ be an enumeration of a dense set in $X$. Since the target $Y$ is a compact metric space, we can find a subsequence $f_1^n = f_{n_k}$ for which $f_{n_k}(x_1)$ converges. We can then find a subsequence $f_2^n$ of the subsequence $f_1^n$ for which $f_2^n(x_2)$ converges. Similarly we define inductively $f_k^n$ to be a subsequence of $f_{k-1}^n$ for which $f_k^n(x_k)$ converges. It follows that the diagonal subsequence $f_n^n$ (the $n$th element of the $n$th subsequence) converges at each point $x_m$.

We claim that for any $\varepsilon > 0$ there exists an $N_\varepsilon$ such that for all $n, m > N_\varepsilon$ and all $x \in X$ we have $d_Y(f^n_n(x), f^n_m(x)) < \varepsilon$. To verify the claim, let $\delta$ be such that if $d_X(x_1, x_2) < \delta$ then $d_Y(f(x_1), f(x_2)) < \varepsilon/3$ for any $f \in F$. Such a $\delta$ exists by our assumption that the family $F$ is equi-continuous. Let $B_i$ be the open ball of radius $\delta$ centered at $x_i$, consisting of points $x$ such that $d_X(x, x_i) < \delta$. Since the set $x_i$ is dense the balls $B_i$ cover $X$. Since $X$ is compact there is a finite set of balls, say $B_1, \ldots, B_l$ which covers $X$. Since the sequences $f^n_m(x_i)$, are convergent there are $N_i$, such that for all $n, m > N_i$, $d_Y(f^n_m(x_i), f^n_m(x_i)) < \varepsilon/3$. Choose $N_\varepsilon$ such that $N_\varepsilon > N_i$ for all $1 \leq i \leq l$. Given $x \in X$ let $x_i, 1 \leq i \leq l$ be such that $d_X(x, x_i) < \delta$. If $n, m > N_\varepsilon$ then by the triangle inequality

$$d_Y(f^n_n(x), f^n_m(x))$$

$$\leq d_Y(f^n_n(x), f^n_n(x_i)) + d_Y(f^n_n(x_i), f^n_m(x_i)) + d_Y(f^n_m(x_i), f^n_m(x))$$

$$< \varepsilon/3 + \varepsilon/3 = \varepsilon$$

We see that for any $x \in X$, the sequence $f^n_n(x)$ is Cauchy and hence, since $Y$ is compact, converges to some value $f(x)$. In addition, for all $x \in X$ and any $n > N_\varepsilon$ we have $d_Y(f_n(x), f(x)) \leq \varepsilon$ by the uniformity of the Cauchy sequence condition. We see that $f^n_n$ converges uniformly to $f$ and a
triangle inequality argument similar to the one above shows that \( f \) will be (uniformly) continuous. \( \quad \text{q.e.d.} \)

2. measure theory

Definition 156. Let \( X \) be a set. A \( \sigma \)-algebra on \( X \) is a non-empty family \( \Sigma \) of subsets of \( X \) such that

(i) If \( A \in \Sigma \) then \( X - A \in \Sigma \).

(ii) If \( A_i \in \Sigma \), \( i = 1, 2, \ldots \), then \( \bigcup_i A_i \in \Sigma \).

The pair \((X, \Sigma)\) is said to be a measurable space, the subsets in \( \Sigma \) are called the measurable sets. If \((X, \Sigma_X)\) and \((Y, \Sigma_Y)\) are a pair of measurable spaces and \( f : X \to Y \) is a function between the sets, we say that \( f \) is measurable if for any \( A \in \Sigma_Y \) we have \( f^{-1}(A) \in \Sigma_X \).

Definition 157. A (non-negative) measure on a measurable space \((X, \Sigma)\) is a function \( \mu \) from \( \Sigma \) to \( \mathbb{R} \cup \{\infty\} \) such that

(i) \( \mu(A) \geq 0 \) for all \( A \in \Sigma \) and \( \mu \) vanishes on the empty set.

(ii) If \( A_i \in \Sigma \), \( i = 1, 2, \ldots \) are pairwise disjoint, then \( \mu(\bigcup_i A_i) = \sum_{i=1}^{\infty} \mu(A_i) \).

We call the triple \((X, \Sigma, \mu)\) a measure space. If \( \mu(X) = 1 \) we say that \( \mu \) is a probability measure.

Definition 158. Let \((X, \Sigma, \mu)\) be a measure space. Let \( T : X \to X \) be a transformation on \( X \). We say that \( T \) is measure preserving if \( T \) is measurable and for any \( A \in \Sigma \) we have \( \mu(T^{-1}(A)) = \mu(A) \). If \( T \) is measure preserving, we call the quadruple \((X, \Sigma, \mu, T)\) a dynamical system. We say that a set \( A \in \Sigma \) is \( T \) invariant, if \( T^{-1}(A) = A \). We say that \( T \) is ergodic if for any \( T \) invariant set we have \( \mu(A) = 0 \) or \( \mu(A) = \mu(X) \).

the following basic result is called the Borel-Cantelli lemma.

Lemma 23. Let \( \mu \) be a measure and \( A_j \) sets such that \( \sum_j \mu(A_j) < \infty \). Let \( B \) be the set of \( x \) which belong to infinitely many of the \( A_j \), i.e.,

\[
B = \bigcap_n (\bigcup_{j=n}^{\infty} A_j)
\]

Then, \( \mu(B) = 0 \).

Proof: We have \( B = \bigcap_n (\bigcup_{j=n}^{\infty} A_j) \subset \bigcup_{j=n}^{\infty} A_j \), but

\[
\mu(\bigcup_{j=n}^{\infty} A_j) \leq \sum_{j=n}^{\infty} \mu(A_j)
\]

and the last sum converges to zero because the series has a finite sum. \( \quad \text{q.e.d.} \)
DEFINITION 159. A set $A$ in a measure space is said to be a null set if $A \subset B$ for some $B \in \Sigma$ such that $\mu(B) = 0$. We say that the measure space is complete if $\Sigma$ contains all the null sets. We say that a $\sigma$-algebra $\Sigma$ is the completion of $\Sigma$ if it is the smallest $\sigma$-algebra on $X$ containing $\Sigma$ and the null sets.

We say that an $X$ dependent property holds almost everywhere (a.e.) if it holds outside a null set.

DEFINITION 160. If $X$ is a topological space, the Borel $\sigma$-algebra is the smallest $\sigma$-algebra containing the open sets of $X$.

DEFINITION 161. Consider $\mathbb{R}^n$ equipped with the topology of the Euclidean metric. A Lebesgue measure $\mu$ on $\mathbb{R}^n$ is a measure on the completion $\Sigma$ of the Borel $\sigma$-algebra which is

(i) Translation invariant, i.e., $\mu(T^{-1}_x(A)) = \mu(A)$ for all $x \in \mathbb{R}^n$ and $A \in \Sigma$, where $T_x(y) = x + y$ is translation by $x$.
(ii) Let $I^n$ be the unit cube, then $\mu(I^n) = 1$.

The following is a basic theorem of Lebesgue, [120].

THEOREM 72. There exists a unique Lebesgue measure on $\mathbb{R}^n$.

DEFINITION 162. Let $\mu$ be the Lebesgue measure on $\mathbb{R}^n$ and let $A$ be a measurable set. For $x \in \mathbb{R}^n$ we define $d_x(A)$, the density of $A$ at $x$, as

$$
\lim_{\varepsilon \to 0} \frac{\mu(A \cap B(x, \varepsilon))}{\mu(B(x, \varepsilon))}
$$

if the limit exists.

The following theorem is known as the Lebesgue density theorem, see [152].

THEOREM 73. Let $A$ be a measurable set w.r.t. the Lebesgue measure. Then, $d_x(A) = 1$ almost everywhere in $A$.

As a consequence, we can prove the following result.

THEOREM 74. Let $t_n$ be a sequence such that $t_n \to 0$. If a measurable function $f$ is $t_n$ periodic for all $n$ then $f$ is an almost everywhere constant function.

Proof: Fix some interval $[a, b]$. We claim that for any $c \in \mathbb{R}$, the set $A_c = \{x \mid f(x) \leq c\} \cap [a, b]$ has Lebesgue measure $\mu$, either 0 or $b - a$. Assume $\mu(A_c) > 0$. By the Lebesgue density theorem, there exists a point $x \in A_c$ such that $d_x(A_c) = 1$. Fix $\varepsilon > 0$ and let $t_n$ be such that $t_n < \varepsilon/4$ and

$$
d(t_n/2, x) = \frac{\mu(A_c \cap B(x, t_n/2))}{\mu(B(x, t_n/2))} \geq \frac{b - a - \varepsilon}{b - a - \varepsilon/2}
$$

Let $k$ be such that $b > x + (k + 1/2)t_n > b - t_n$ and let $l$ be such that $a < x - (l + 1/2)t_n < a + t_n$. Since $f$ is periodic of period $t_n$ so is $A_c$. The
interval \([x-(l+1/2)t_n, x+(k+1/2)t_n]\) consists of a disjoint (except for edge points) union of translates by multiples of \(t_n\) of the interval \([x-t_n/2, x+t_n/2]\) = \(B(x, t_n/2)\) and hence the density

\[
\frac{\mu([x-(l+1/2)t_n, x+(k+1/2)t_n])}{\mu([x-(l+1/2)t_n, x+(k+1/2)t_n])} = d(t_n/2, x)
\]

We also have that

\[
\mu([x-(l+1/2)t_n, x+(k+1/2)t_n]) = x + (k + 1/2)t_n - (x - (l + 1/2)t_n) \geq b - a - 2t_n \geq b - a - \frac{\varepsilon}{2}
\]

We conclude that

\[
\mu(A_c \cap [a, b]) \geq \mu(A_c \cap [x-(l+1/2)t_n, x+(k+1/2)t_n]) \geq (b-a-\varepsilon/2) = b-a-\varepsilon
\]

Since this is true for all \(\varepsilon > 0\) we have \(\mu(A_c \cap [a, b]) = b-a\).

Let \(c_0\) be the supremum of the set \(\{c \mid \mu(A_c \cap [a, b]) = 0\}\). We claim that on the interval \([a, b]\) we have \(f(x) = c_0\) almost everywhere. Consider the set \(A_{<c_0} = \{x \mid f(x) < c_0\} \cap [a, b]\), then \(A_{<c_0} = \bigcup_i A_{c_0-1/i}\), but, by definition \(\mu(A_{c_0-1/i}) = 0\) so we conclude by countable additivity of the measure that \(\mu(A_{<c_0}) = 0\). Similarly, we know that \(A_{>c_0} = \{x \mid f(x) > c_0\} \cap [a, b] = \bigcup_i [a, b] - A_{c_0+1/i}\) and that the latter sets, by the definition of \(c_0\) have measure zero, hence \(f(x) = c_0\) almost everywhere as asserted. Taking an increasing sequence of intervals (by containment), whose union is the entire line we obtain the desired result. \(q.e.d.\)

**Definition 163.** We say that a function \(f\) on a closed interval \(I\) is **absolutely continuous** if for any \(\varepsilon > 0\), there exists a \(\delta > 0\) such that for any finite family of disjoint segments \((a_i, b_i)\), with \(\sum_i b_i - a_i < \delta\) we have \(\sum_i |f(a_i) - f(b_i)| < \varepsilon\).

It is obvious that a Lifschitz function with some constant \(c\) is absolutely continuous, this is seen by taking \(\delta = \varepsilon/c\). The following basic result, whose proof can be found in [152] provides the proper generalization of the fundamental theorem of calculus in the context of Lebesgue integration.

**Theorem 75.** Let \(f\) be an absolutely continuous function from on an interval \(I = [a, b]\), then the derivative \(f'\) exists almost everywhere, is integrable and satisfies

\[
\int_a^b f'(s) ds = f(b) - f(a)
\]

In chapter 2 we use the following theorem, see [152] for a proof.

**Theorem 76.** For any measurable and bounded function \(g\) on \(I\) we can find a continuous function \(f_n\) with the same bound such that \(\int_a^b |f_n - g|^2 ds < 1/n\).
3. Probability

We recall some basic definitions from probability theory.

**Definition 164.** A real valued random variable is a measurable function $X$ from a probability space $(\Omega, \Sigma, \mu)$ to $(\mathbb{R}, B)$ where $B$ denotes the Borel $\sigma$-algebra on $\mathbb{R}$. A finite random variable $X$ is a random variable that has only a finite set of values $a_1, \ldots, a_n$ which are obtained with probabilities $p_1, \ldots, p_n$ respectively. A random variable is discrete if it has a sequence of real values $a_1, a_2, \ldots$ which are obtained with probabilities $p_1, p_2, \ldots$ respectively. Given a real valued random variable $X$ we define the associated distribution function $X(t)$ by the formula

$$X(t) = \Pr(X \leq t) = \mu(X^{-1}([-\infty, t]))$$

A function $f_X(t)$ is called the density of $X$ if $X(t) = \int_{-\infty}^{t} f(s)ds$. Random variables $X_1, \ldots, X_k$ are said to be independent if $\Pr(X_1 \leq a_1, X_2 \leq a_2, \ldots, X_k \leq a_k) = \prod_{j=1}^{k} \Pr(X_j \leq a_j)$. We say that a sequence $X_i$ of random variables are independent and identically distributed (i.i.d.) if for any finite subset of indices $i_1, \ldots, i_k$ the variables $X_{i_j}$ are independent and all the variables $X_{i_j}$ share the same distribution function.

We assign to random variables various quantities.

**Definition 165.** If $X$ is a discrete random variable we define the expectation of $X$, denoted by $E(X)$, to be $E(X) = \sum_{i} p_i a_i$. If $X$ is real valued with density $f_X(t)$ we define $E(X) = \int_{-\infty}^{\infty} t f_X(t)dt$. We define the $k$'th moment to be $E(X^k)$. The variance $V(X)$ of a random variable is given by $V(X) = E((X - E(X))^2) = E(X^2) - E(X)^2$. The standard deviation $\sigma(X)$ is given by $\sigma(X) = \sqrt{V(X)}$.

By definition, if $X, Y$ are any random variables then $E(X+Y) = E(X) + E(Y)$. If random variables $X, Y$ are independent then it follows from the definition that $E(XY) = E(X)E(Y)$ and after a short computation we see that $V(X+Y) = V(X) + V(Y)$. Consequently, if $X_1, \ldots, X_k$ are i.i.d., then, $E(\sum_{i} X_i) = kE(X)$, where $X$ refers to a generic random variable with the common distribution function $X(t)$ and $V(\sum_{i} X_i) = kV(X)$ or equivalently $\sigma(\sum_{i} X_i) = \sqrt{k}\sigma(X)$.

The following nearly trivial estimate is known as Markov’s inequality.

**Lemma 24.** Let $X \geq 0$ be a random variable with non-negative values. Let $s \geq 0$, then

$$E(X) \geq \Pr(X \geq s)s$$

**Proof:** The right hand side is the expectation of the random variable $Y_s$ which takes the value 0 when $X < s$ and takes the value $s$ when $X \geq s$. Since $X(t) \geq Y_s(t)$ for all $t$ we obtain the inequality. q.e.d.

The next result, which is an immediate consequence of Markov’s inequality is known as Chebyshev’s inequality.
Lemma 25. The following estimate holds for any \( s \geq 0 \).

\[
Pr(|X - E(X)| \geq \sqrt{s}) \leq \frac{V(X)}{s}
\]

**Proof:** Applying Markov’s inequality to the random variable

\[
Z = (X - E(X))^2 = E(X^2) - (E(X))^2
\]

we obtain the inequality \( V(X) = E(Z) \geq Pr((X - E(X))^2 \geq s)s = Pr(|X - E(X)| \geq \sqrt{s})s \) which is Chebyshev’s inequality. \( \text{q.e.d.} \)

From Chebyshev’s inequality we easily obtain the weak law of large numbers.

**Theorem 77.** Let \( X_i, i = 1, ..., h \) be a sequence of random variables, then

\[
(203) \quad Pr(\left| \sum X_i - hE(X) \right| \geq c\sigma(X)\sqrt{h} \leq \frac{1}{c^2})
\]

**Proof:** It is trivial to verify that for a random variable \( X \) and a constant \( c \) we have \( V(X - c) = V(X) \) and \( V(cX) = c^2V(X) \). Let

\[
Y_h = \frac{\sum_{i=1}^{h} X_i - hE(X)}{\sigma(X)\sqrt{h}}
\]

then \( E(Y) = 0 \) and \( V(Y_h) = \frac{hV(X)}{hV(X)} = 1 \). We conclude from Chebyshev’s inequality that

\[
(204) \quad Pr(|Y_h| \geq c) \leq \frac{1}{c^2}
\]

for all \( c > 0 \) and for all \( h \), and this is equivalent to (203). \( \text{q.e.d.} \)

**Definition 166.** Given a discrete probability distribution \( Z = (p_i) \) we define the **generating function** \( f(x) \) associated with \( p_i \) by the formula

\[
f(x) = f_Z(x) = \sum_{i=0}^{\infty} p_i x^i.
\]

Generating functions are a convenient and useful way of summarizing the information present in a distribution \( Z \). In particular the moments of \( Z \) can be computed from the derivatives of \( f \) evaluated at \( x = 1 \), the first three relations being

\[
(205) \quad f(1) = \sum_i p_i = 1
\]

\[
(206) \quad f'(1) = \sum_i ip_i = E(Z_1)
\]

\[
(207) \quad f''(1) = \sum_i i(i - 1)p_i = \sum_i (i^2 - i)p_i = E(Z_1^2) - E(Z_1)
\]

A simple inspection of the way that power series are multiplied shows that if \( Y, Z \) are independent random variables then \( f_{Y+Z} = f_Y f_Z \).
3.1. Basic Markov chain theory. We first recall some basic definitions and results about discrete Markov chains.

**Definition 167.** Consider a sequence of random variables $X_i$, which all take values in the non-negative integers. We call the values, the states of the process. We say that the sequence $X_i$ forms a Markov chain, if the value of $X_{i+1}$ depends on the value of $X_i$, but not on the values of $X_j$, $j < i$. Thinking of $i$ as representing the present time, we may say that the future depends on the present but not on the past.

We let $p_{m,l}^i = \Pr(X_{i+1} = l \mid X_i = m)$ be the probability of transitioning from state $m$ to state $l$ at time $i$.

**Definition 168.** If $p_{m,l}^i$ is independent of $i$ for all $m,l$, then we say that the Markov chain is homogeneous or stationary.

We will only consider stationary Markov chains. We denote the stationary (time independent) transition probabilities by $p_{m,l}$.

**Definition 169.** We say that the Markov chain is irreducible if for any pair of states $m,l$, and given that, $X_i = m$, there exists a $j = j(m,l)$ such that $\Pr(X_{i+j} = l) > 0$, i.e., there is a positive probability of transitioning from the present state $m$ to the state $l$, sometime in the future (in $j$ time steps).

**Definition 170.** A state $m$ is called periodic if the set of times $j$ such that $\Pr(X_{i+j} = m \mid X_i = m) > 0$ is an arithmetic sequence with difference greater than 1. A state is called aperiodic if it is not periodic.

We further classify states by looking more closely at their return probabilities.

**Definition 171.** Let $f_m$ denote the probability that starting in state $m$, i.e., $X_1 = m$, the process will ever return to the state $m$. We say that a state $m$ is persistent if $f_m = 1$. Otherwise, we say that the state is transient.

Persistent states are further classified.

**Definition 172.** Let $r_j^m$, be the probability that the process with $X_1 = m$ will return to state $m$ for the first time after exactly $j$ time steps. If $j$ is the first time we returned to $m$ we say that $j$ is the return time of $m$. We let $\mu_m = \sum_j j r_j^m$ denote the average return time of state $m$. We say that a persistent state is a null state if $\mu_m = \infty$, otherwise we say that the state is ergodic.

Consider, for example, the set of transition probabilities $p_{m,m-1} = 1$ for $m$ odd and $p_{m,m+1} = 1$ for $m$ even. In this case each state returns to itself after an even number of steps, but never after an odd number of steps, hence it is periodic. We also note that this example is obviously not irreducible.

As another example, consider the chain with $p_{m,m} = 1/2$ and $p_{m,m+1} = 1/2$. With probability $1/2$ a chain starting from $m$ will transition to a state
with higher index and will never come back to state \( m \). In this case the states are transient.

In an irreducible chain, it is easy to see that all states have the same classification, either all states are transient or all are persistent and null or all are ergodic. In addition, either all states are periodic or all of them are aperiodic. We conclude that for an irreducible chain, we can apply these terms to the entire chain unambiguously.

Given a Markov chain with some initial distribution, \( X_1 \) and transition probabilities \( p_{m,l} \), we are interested in the sequence of probabilities \( p_i^l = \text{Pr}(X_i = l) \). We are even more interested in the asymptotic behavior of these numbers as \( i \to \infty \).

**Definition 173.** We say that the chain has a **stationary distribution** \( p_l \) if for all possible distributions \( X_1 \), we have \( \lim_{i \to \infty} p_i^l = p_l \), for all \( l \), and \( \sum_l p_l = 1 \). This means that the probability of being in state \( l \) after many time steps converges to \( p_l \) independently of the initial condition of the system.

The stationary distribution, if it is to exist, must satisfy a certain consistency condition. Given the transition probabilities \( p_{m,l} \) we have \( p_i^{l+1} = \sum_m p_m^i p_{m,l} \), which simply means that we consider all the possible states at time \( i \) according to their probabilities and consider the possibility of transitioning from them to \( l \). Taking limits from both sides of the equation we see that we must have

\[
(208) \quad p_l = \sum_m p_m p_{m,l}
\]

We have the following basic result in the theory of Markov chains, sometimes known as Kac’s lemma, a proof can be found in [105] or in [72] volume II.

**Theorem 78.** Assume that \( X_i \) is an irreducible and aperiodic Markov chain. Assume that the chain is ergodic. Then, there exists a stationary distribution \( p_l \) and \( p_l = 1/\mu_l \), where \( \mu_l \) is the mean return time of state \( l \). Conversely, if there are numbers \( p_l \geq 0 \) which satisfy (208) and \( \sum_l p_l = 1 \), the chain is ergodic and the \( p_l \) form the stationary distribution.

### 4. Algebra

#### 4.1. Groups.

**Definition 174.** Let \( (\mathbb{Z}_N)^* \) denote the group of integers \( 0 \leq a \leq N - 1 \) such that \( \gcd(a, N) = 1 \), with multiplication given modulo \( N \). A **(Dirichlet) character modulo** \( N \) is a multiplicative homomorphism \( \chi : (\mathbb{Z}_N)^* \to \mathbb{C}^* \), where \( \mathbb{C}^* \) denotes the group of non-zero complex numbers with multiplication.

Since \( (\mathbb{Z}_N)^* \) is finite, the image of a Dirichlet character consists of roots of unity. Given a character \( \chi \) we may compose it with complex conjugation, which on roots of unity (or the unit circle) is the group isomorphism sending \( g \) to \( g^{-1} \). This produces a character which we denote by \( \bar{\chi} \). The characters
form a group via function multiplication and \( \chi = \chi^{-1} \). The trivial character which sends everything to 1 is obviously the unit.

Consider some \( M | N \). There is an obvious projection homomorphism \( \pi_{N,M} : (\mathbb{Z}_N)^* \to (\mathbb{Z}_M)^* \) which maps \( a \ (\mod N) \) to \( a \ (\mod M) \). Let \( \chi_M : (\mathbb{Z}_M)^* \to \mathbb{C}^* \) be a Dirichlet character modulo \( M \), then we can use \( \chi_M \) to construct a Dirichlet character \( \chi_N \) modulo \( N \) as the composition of homomorphisms \( \chi_N = \chi_M \pi_{N,M} \).

**Definition 175.** We say that a Dirichlet character modulo \( N \) is **primitive** if it does not arise this way from a Dirichlet character \( \chi_M \) with \( M | N \). If a Dirichlet character \( \chi \) modulo \( N \) is primitive we say that \( N \) is the **conductor** of \( \chi \).

We may extend a Dirichlet character to a function \( \chi : \mathbb{Z} \to \mathbb{C} \) by \( \chi(n) = \chi_N(n \ (\mod N)) \), with the understanding that if \( \gcd(n,N) > 1 \) then \( \chi(n) = 0 \). This mapping is still multiplicative, namely \( \chi(nm) = \chi(n)\chi(m) \).

Let \( q \) be a prime. Let \( \mathbb{F}_q \) be the field with \( q \) elements, consisting of residue classes modulo \( q \). Let \( \mathbb{F}_q^* \) denote the multiplicative group of non zero elements of the field.

**Lemma 26.** The group \( \mathbb{F}_q^* \) is cyclic

**Proof:** This group is obviously commutative. In addition, for each integer \( k > 0 \) it contains at most \( k \) elements \( x \) such that \( x^k = 1 \), since a degree \( k \) polynomial in a field has at most \( k \) roots. The fundamental structure theorem for finite commutative groups says that any such group is isomorphic to a direct sum of cyclic groups \( G_i \), \( i = 1, ..., n \), whose orders \( m_i \) satisfy \( m_i | m_{i+1} \). It is obvious however that if there is more than one cyclic group in the decomposition, i.e., \( n > 1 \), then we will have \( m_1^2 \) solutions to the order equation \( x^{m_1} = 1 \). We conclude that \( n = 1 \) or stated otherwise, \( \mathbb{F}_q^* \) is a cyclic group. We note that this argument works for the multiplicative group of any finite field with the same conclusion. q.e.d.

**5. Linear algebra**

**Definition 176.** Let \( M \) be a real \( n \) by \( n \) matrix. The **bilinear form** associated with \( M \) is the function \( B_M : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) given by \( B_M(v,w) = vMw^t \). The **quadratic form** associated with \( M \) is the function \( Q_M : \mathbb{R}^n \to \mathbb{R} \) given by \( Q_M(v) = B_M(v,v) = vMv^t \).

**Definition 177.** Let \( K \) be a \( d \) by \( d \) symmetric non-singular matrix all of whose eigenvalues are positive. We call such a matrix a **Euclidean matrix.** The **Euclidean metric** associated with \( K \) is the function \( d_K(v) = \sqrt{vKv^t} \).

The standard example is given by the identity matrix \( I \), which through the expression \( vIw^t \) leads to the usual notions of Euclidean length and orthogonality. As with the case of Minkowski metrics, Sylvester’s theorem states that all Euclidean metrics \( K \) are of the form \( LIL^t = LL^t \), where \( L \) is an invertible transformation. Since the matrix \( K \) is symmetric it has a basis
of orthogonal (w.r.t. $I$) eigenvectors. From this and the positivity of the eigenvalues we can easily deduce that for any vector $v$ the expression $vKv^t$ is always non-negative and hence that the definition of a Euclidean metric makes sense. Taking an orthogonal basis also shows that for any vector $v$ we have $\sqrt{\lambda_{\min}} ||v|| \leq ||v||_K \leq \sqrt{\lambda_{\max}} ||v||$ where $\lambda_{\min}, \lambda_{\max}$ are respectively, the minimal and maximal eigenvalues of $K$.

6. Harmonic analysis and transforms

We consider several transforms and some of their basic properties.

**Definition 178.** Let $g$ be a continuous complex valued function on the real line such that $\lim_{x \to -\infty} |g(x)x^n| = 0$ and $\lim_{x \to \infty} |g(x)x^n| = 0$ for all integer $n$. For such functions we define the Fourier transform $\hat{g}$ by the formula

$$\hat{g}(x) = \int_{\mathbb{R}} g(y)e^{-2\pi ixy}dy$$

The following are well known basic results concerning the Fourier transform.

**Theorem 79.** The Fourier transform satisfies the equality

$$\int_{\mathbb{R}} |g^2(x)|dx = \int_{\mathbb{R}} |\hat{g}^2(x)|dx$$

This equality is known as the Plancherel formula.

**Theorem 80.** Let $f(x) = e^{-\pi x^2}$, then $\tilde{f} = f$.

**Proof:** We differentiate $\tilde{f}(y)$ under the integral sign and use integration by parts

$$\frac{d}{dy} \tilde{f} = -2\pi i \int_{\mathbb{R}} e^{-2\pi xy}e^{-\pi x^2}dx$$

$$= -2\pi y \int_{\mathbb{R}} e^{-2\pi ixy}f(x) = -2\pi y \tilde{f}$$

or

$$\frac{d}{dy} (\log(\tilde{f})) = -2\pi y$$

which has the solution $\tilde{f}(y) = ce^{-\pi y^2}$. By the Plancherel formula $c$ is either 1 or $-1$ and checking $y = 0$ shows that $c$ is sign preserving and hence equals 1. q.e.d.

**Theorem 81.** For any continuous exponentially decaying function $f$ we have

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \tilde{f}(n)$$

This equality is known as the Poisson summation formula.
Proof: We define \( g(x) = \sum_{n=-\infty}^{\infty} f(x+n) \). The function \( g(x) \) is periodic and infinitely differentiable. Let \( \sum_{n} c_{n} e^{2\pi inx} \) be its Fourier expansion. Since \( g \) is differentiable, its Fourier expansion is known to converge to \( g \) at all points, [106]. By definition we have

\[
c_{n} = \int_{0}^{1} g(x) e^{-2\pi inx} \, dx = \int_{0}^{1} \sum_{n=-\infty}^{\infty} f(x+n) e^{-2\pi inx} \, dx
\]

Substituting \( x = 0 \) we obtain

\[
g(0) = \sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} c_{n} = \sum_{n=-\infty}^{\infty} \tilde{f}(n)
\]
as required q.e.d.

Consider the unit circle \( S^1 \) in the complex plane consisting of \( z \in \mathbb{C} \) with \( |z| = 1 \). This set forms a group w.r.t. multiplication. We can parametrize the points of the group by elements \( x \in [0,1] \) by mapping \( x \) to the point \( z \) at angle \( 2\pi x \). With this parametrization, a continuous function on \( S^1 \) is the same as a continuous periodic function with period 1 on \( \mathbb{R} \).

We note that the continuous homomorphisms from \( \mathbb{R} \) to \( S^1 \) are precisely the functions \( e^{2\pi ixy} \). Such homomorphisms are called unitary characters. Any such homomorphism is an eigenvector of the linear operator \( T_{x}(f)(y) = f(x+y) \), which acts on functions by shifting the variable by a constant group element. When the group is commutative, all the shift operators commute. The Fourier transform is obtained by integrating a function against all the complex conjugates of these homomorphisms w.r.t. an invariant measure (the Lebesgue measure in the case of \( \mathbb{R} \)) w.r.t. the group translations. This is the analogue in infinite dimensional spaces of functions) of the idea of writing a general vector as a combination of eigenvectors (for the shift operators). This general setup can be carried for groups other than \( \mathbb{R} \), resulting in more Fourier transforms. When the group is \( S^1 \), the continuous homomorphisms of \( S^1 \) to \( S^1 \) are of the form \( e^{2\pi inx} \), \( n \in \mathbb{Z} \). The \( n \)th fourier coefficient of a function \( f \) will then be \( \int_{0}^{1} f(x) e^{-2\pi inx} \, dx \).

In chapter 2 we will need the following lemma which states that we can approximate any continuous function on \( S^1 \) by a finite combination of characters.

Lemma 27. Fix \( \varepsilon > 0 \). For any continuous function \( f(x) \) with \( f(x) = f(x+1) \) there is a polynomial \( \phi(x) \) in the functions \( e^{2\pi ix}, e^{-2\pi ix} \) such that for all \( x \) we have \( |f(x) - \phi(x)| < \varepsilon \).

Proof: We consider the functions \( g_{k}(x) = A_{k}(\frac{1+\cos(x)}{2})^{k} \) where \( A_{k} \) is chosen so that \( \int_{0}^{1} g_{k}(x) \, dx = 1 \). To estimate \( A_{k} \) we need to estimate
\[ f_0 \left( \frac{1 + \cos(x)}{2} \right)^k dx. \] Using the Taylor expansion we get \( \frac{1 + \cos(x)}{2} = 1 - x^2/4 + o(x^3) \) for \( x \) small and hence for \( k \) large enough and \( |x| \leq 1/\sqrt{k} \) we have \( \left( \frac{1 + \cos(x)}{2} \right)^k \geq (1 - \frac{1}{4k})^k > e^{-1/2} \) and so \( \int_0^1 \left( \frac{1 + \cos(x)}{2} \right)^k dx > \frac{2e^{-1/2}}{\sqrt{k}} \). We conclude that \( A_k < C\sqrt{k} \) for some constant \( C > 0 \). We note that for any given \( 1/4 > \eta > 0 \) we have for \( \eta < x < 1 - \eta \) the inequality
\[
\left( \frac{1 + \cos(x)}{2} \right) \leq (1 + \cos(\eta))/2 = \eta < 1
\]
which implies that
\[
A_k \int_{\eta}^{1-\eta} \left( \frac{1 + \cos(x)}{2} \right)^k dx \leq C\sqrt{k}\eta^k \to 0
\]
Since \( \cos(x) = \frac{1}{2}(e^{2\pi ix} + e^{-2\pi ix}) \) it is clear that \( g_k(x) \) is a polynomial in \( e^{2\pi ix}, e^{-2\pi ix} \). Continuous functions with period 1 carry an operation known as a convolution. It is defined via the formula
\[
f \ast g(x) = \int_0^1 f(y)g(x-y)dy
\]
The convolution operation is commutative as can be seen by making the change of variables \( y \to x-y \). Let \( k_m(x) = e^{2\pi mx} \), then
\[
f \ast k(x) = \int_0^1 f(y)e^{2\pi m(x-y)}dy = (\int_0^1 f(y)e^{-2\pi y}dy)e^{2\pi mx} = (\int_0^1 f(y)e^{-2\pi y}dy)k_m
\]
Consequently, for any polynomial \( \phi(x) \) in \( e^{2\pi ix}, e^{-2\pi ix} \) the function \( f \ast \phi \) is also a polynomial in \( e^{2\pi ix}, e^{-2\pi ix} \), which we can write as
\[
\phi(x) = \sum_{m=-l}^l a_ke^{2\pi mx}
\]
Since \( \int_0^1 e^{2\pi mx} dx = 0 \) for all integer \( m \neq 0 \), we have \( \int_0^1 \phi(x)dx = a_0 \) and also \( \int_0^1 f \ast \phi(x)dx = a_0(\int_0^1 f(x)dx) \). Applying to \( g_k \) we get \( \int_0^1 f \ast g_k(x)dx = \int_0^1 f(x)dx \).

Since \( f \) is a continuous function on a closed interval, \( F \) is uniformly continuous, hence for any \( \delta > 0 \) there is an \( \eta > 0 \) such that \( |x - y| < \eta \) implies \( |f(x) - f(y)| < \delta \). Also since \( f \) is continuous it is bounded, say \( |f(x)| < M \). We have the estimate
\[
|g_k(x) \ast f(x) - f(x)| = |A_k \int_0^1 (f(y) - f(x))g_k(x-y)dy|
\]
\[
\leq A_k \int_0^1 |f(y) - f(x)||g_k(x-y)|dy
\]
\[
= A_k \int_{|x-y|<\eta} |f(y) - f(x)||g_k(x-y)|dy + A_k \int_{|x-y|\geq\eta} |f(y) - f(x)||g_k(x-y)|dy
\]
\[ \leq \delta + 2MC\sqrt{kq} \eta < \varepsilon \]

the last inequality for \( \delta \) small enough and \( k \) large enough and so \( g_k \ast f \) is the required polynomial. \( \text{q.e.d.} \)

**Definition 179.** For a density \( f(t) \), we define the Laplace transform for \( s \geq 0 \) by the formula

\[
B(s) = \int_0^\infty e^{-st} f(t) dt
\]

Assuming, for example, that \( f \) is bounded, the transform will exist. This transform is essentially the Fourier transform for functions on the real line computed for the argument \( is \), where we take \( f(t) = 0 \) for \( t < 0 \). One reason for the interest in the Laplace transform in probability stems from the fact that it allows us to compute moments of \( X \) through successive differentiation. We have \( E(X^n) = (-1)^n B^{(n)}(0) \), where \( B^{(n)} \) refers to the \( n \)’th derivative of the transform. To see this we differentiate \( B \) under the integral sign, assuming again that \( f(t) \) is nice enough to make this procedure legal

\[
B^{(n)}(s) = \int_0^\infty (-t)^n e^{-st} f(t) dt
\]

and setting \( s = 0 \) gives the desired conclusion.

7. Complex analysis

We recall some basic definitions.

**Definition 180.** A complex valued function on a domain in \( \mathbb{C} \), is said to be holomorphic at a point \( z_0 \) if it is differentiable w.r.t. \( z \) as a complex variable, or equivalently, if \( f(z) \) can be written in a neighborhood of \( z_0 \) in the form of a convergent power series \( f(z) = \sum_{n=0}^{\infty} a_n z^n \). We say that a function is (complex) analytic (or holomorphic) if it is holomorphic at every point of the domain. We say that \( f(z) \) is meromorphic at \( z_0 \) if for some neighborhood \( U \) of \( z_0 \), \( f(z) \) can be written as a convergent Laurent series \( f(z) = \sum_{n=-m}^{\infty} a_n z^n \) on \( U \setminus \{z_0\} \), for some \( m \geq 0 \).

8. Basic definitions in graph theory

We review some very basic definitions in graph theory, readers who have basic familiarity with the subject can skip to the next section.

**Definition 181.** A graph \( G \) consists of a set \( V = V(G) \) called the vertices and a set \( E = E(G) \) of unordered pairs of the form \( \{u, w\} \) where \( u, w \in V \), called the edges. Given an edge \( e = \{u, w\} \), we say that \( u \) and \( v \) are incident to \( e \). We also say that the edge \( e \) connects \( u \) and \( v \). For simplicity we will denote edges by \( (u, v) \) rather than \( \{u, w\} \). The degree of a vertex is the number of edges which are incident to it. We say that a graph \( G \) is \( d \)-regular if each vertex \( v \in V(G) \) is incident to exactly \( d \) edges.
Given two disjoint subsets $A, B \subset V(G)$ we let $E(A, B)$ denote the set of all edges that connect a vertex of $A$ with a vertex of $B$.

The number of vertices of a graph will be denoted by $|V|$ and the number of edges by $|E|$. We note that in a $d$ regular graph $|E| = d|V|/2$.

**Definition 182.** Given a graph $G$ we define the adjacency matrix of $G$ to be the $|V| \times |V|$ matrix $A$ whose rows and columns are indexed by vertices of $G$. If $u, v \in V(G)$ then $a_{uv} = 1$ if $(u, v) \in E$ and 0 otherwise.

**Definition 183.** A (non backtracking) walk (or path) $W$ in $G$ from a vertex $v$ to a vertex $u$ is a sequence of vertices $v = v_0, v_1, \ldots, v_m = u$ in $G$ such that $(v_i, v_{i+1}) \in E(G)$ and $v_i \neq v_{i+2}$ for all $i = 0, \ldots, m - 2$. The length $l(W)$ of such a walk is $m$.

The distance between two sets of vertices $A$ and $B$ in a graph, is the length of the shortest walk from a vertex $a \in A$ to a vertex $b \in B$.

A graph $G$ is connected if any two vertices in $G$ can be connected by a walk. A connected component of a graph $G$ is a maximal (with respect to vertex inclusion) connected subgraph of $G$.

A cycle in $G$ is a connected 2-regular subgraph of $G$. The length of a cycle $C$, denoted $l(C)$, is the number of its edges, which also equals the number of its vertices.

The girth of a graph $G$, denoted $g(G)$, is the minimal cycle length in $G$.

We note that if $C$ is a cycle and $v$ is a vertex of $C$, then $C$ induces two walks of length $l(C)$ from $v$ to itself by going around the cycle in either direction, hence the number of walks of length $k$ from $v$ to itself is at least twice the number of cycles of length $k$ containing $v$.

**Definition 184.** A tree is a connected graph with no cycles.

A leaf is a vertex of a tree with degree 1. A unicycle is a connected graph with a unique cycle.

A rooted tree is a tree with a distinguished vertex $v$.

An acyclic graph (or a forest) is a graph without cycles, or in other words, a graph all of whose connected components are trees.

There is a unique walk between any two vertices of the tree, since the union of two distinct walks will lead to a cycle after the elimination of edges which appear in both walks. It is easy to show by induction that a connected graph is a tree if and only if $|V| = |E| + 1$. We see that a tree always has leaves since $|E|/|V| = (\sum_v \deg(v))/(2|V|) < 1$ and hence there is a vertex whose degree satisfies $\deg(v) < 2$, hence a leaf.

A connected graph is a unicycle if and only if $|V| = |E|$, since removing any edge from the unique cycle will lead to a tree, while adding any edge to a tree will close a unique cycle together with the walk in the tree between the vertices which are incident to the added edge.

**Definition 185.** Given a graph $G$ an orientation of $G$ is a function $f : E \to V$, such that for all $e \in E(G)$, $f(e)$ is incident to $e$. 
A rooted tree has a canonical orientation which is constructed as follows. Let \( r \) be the root and let \( e = (u, v) \) be an edge of the tree. Assume that the unique walk from \( r \) to \( v \) passes through \( u \), then we let \( f(e) = v \). Otherwise the unique walk from \( r \) to \( u \) is the concatenation of the unique walk from \( r \) to \( v \) together with the edge \( e \). In this case we let \( f(e) = u \). This orientation, orients the edges away from the root. We note that the root is not in the image of \( f \). The canonical orientation of a rooted tree is a 1-1 function since otherwise we would have two distinct walks from the root to the same vertex.

**Definition 186.** Given two graphs \( G_1, G_2 \). A **graph homomorphism** \( F : G_1 \rightarrow G_2 \) consists of a map \( F_V : V(G_1) \rightarrow V(G_2) \), such that for each edge \((v, w) \in E(G_1)\) we have \((F_V(v), F_V(w)) \in E(G_2)\). A graph homomorphism is an **isomorphism** if \( F_v \) is 1-1 and onto, and the inverse \( F^{-1}_V \) is also a graph homomorphism. An **automorphism** is an isomorphism from a graph to itself. we say that a graph \( G \) is **vertex transitive** if for any two vertices \( v, w \in V(G) \) there is an automorphism \( F_{v,w} \) of \( G \) which maps \( v \) to \( w \).

**Definition 187.** In a group \( G \), we say that a set of elements \( g_1, \ldots, g_k \) is **symmetric** if for each \( g_i \) there is a \( g_j \) such that \( g_j = g_i^{-1} \). Given \( G \) and a symmetric set \( S \) as above we may construct a graph \( G_S \) with nodes corresponding to elements \( g \in G \) and edges given by \((g, gg_i)\), \( g_i \in S \). Such graphs are called **Cayley graphs**.

The graph will be connected if the set \( S \) is a set of generators. Also note that the maps \( F_{g'} : g \rightarrow g'g \) induce graph automorphisms and that these automorphisms are vertex transitive.
CHAPTER 7

Appendix B: Proofs of theorems

Proof of theorem 10 and completion of the proof of Stirling’s formula:

Let \( c \) be such that \( Rh + c\sqrt{LRh} \) is an integer.

We begin by counting the number of addresses which contain exactly \( Rh + c\sqrt{LRh} \) bits which are 1.

We will use the following short hand notation

\[
\gamma = \gamma_{c,R} = c\sqrt{LRh}
\]

Applying the weak form of Stirling’s formula \( k! \sim C\sqrt{k}(k/e)^k \) to all the factorials in the binomial coefficient we have as \( h \to \infty \)

\[
B(h, Rh + c\sqrt{LRh}) = B(h, Rh + \gamma) \sim \frac{h^h}{\sqrt{CLRbh} (Lh - \gamma)^{Lh-\gamma}(Rh + \gamma)^{Rh+\gamma}}
\]

We can write \( Lh - \gamma = \frac{Lh-\gamma}{Lh} Lh \) and note that \( \frac{Lh-\gamma}{Lh} = (1 - \frac{c\sqrt{R/L}}{\sqrt{h}}) \) to obtain

\[
(Lh - \gamma)^{Lh-\gamma} = (1 - \frac{c\sqrt{R/L}}{\sqrt{h}})^{Lh}(1 - \frac{c\sqrt{R/L}}{\sqrt{h}})^{-\gamma}(L^h)^h h^{Lh}(Lh)^{-\gamma}
\]

Doing a similar trick for \((Rh + \gamma)^{Rh+\gamma}\) and noting that \( L^R R^L = 2^{-H(L)} \) we get

\[
(Lh - \gamma)^{Lh-\gamma}(Rh + \gamma)^{Rh+\gamma} = (\frac{R}{L})^{2H(L)h}(1 - \frac{c\sqrt{R/L}}{\sqrt{h}})^{Lh}(1 - \frac{c\sqrt{R/L}}{\sqrt{h}})^{-\gamma}(1 + \frac{c\sqrt{L/R}}{\sqrt{h}})^{Rh}(1 + \frac{c\sqrt{L/R}}{\sqrt{h}})^{-\gamma}
\]

we observe that for any \( A \) we have \( (1 - \frac{A}{\sqrt{h}})^{-1} = (1 + \frac{A}{\sqrt{h}})(1 - \frac{A^2}{h})^{-1} \) and hence \( (1 - \frac{A}{\sqrt{h}})^{-\gamma} \sim (1 + \frac{A}{\sqrt{h}})^{\gamma} \). We conclude that

\[
(1 - \frac{c\sqrt{R/L}}{\sqrt{h}})^{-\gamma}(1 + \frac{c\sqrt{L/R}}{\sqrt{h}})^{\gamma}
\]

\[
\sim [(1 + \frac{c\sqrt{L/R}}{\sqrt{h}})(1 + \frac{c\sqrt{R/L}}{\sqrt{h}})]^{\gamma}
\]

\[
\sim [1 + \frac{c(\sqrt{L/R} + \sqrt{R/L})}{\sqrt{h}}]^{\gamma}
\]

\[
\sim (1 + \frac{c^2}{\gamma})^{\gamma} \sim e^{c^2}
\]
Writing $m_R = \frac{\sqrt{h}}{c\sqrt{L/R}}$ and $m_L = \frac{\sqrt{h}}{c\sqrt{R/L}}$ we can write
\[
(1 + \frac{c\sqrt{L/R}}{\sqrt{h}})Rh(1 - \frac{c\sqrt{R/L}}{\sqrt{h}})^Lh
= [(1 + \frac{1}{m_R})^{m_R}(1 - \frac{1}{m_L})^{m_L}]^\gamma
\]
Taking an integer $m$ we can estimate the difference $e - (1 + \frac{1}{m})^m$ by considering the Taylor series for $e$ and the binomial expansion of $(1 + \frac{1}{m})^m$, considering only the coefficient of $1/m$ and ignoring higher powers such as $1/m^2$. The $k + 1$ summand of the expansion will have the form
\[
\frac{k(k+1)}{2} \frac{1}{m^k} = \frac{1}{(k+1)!} - \frac{k}{2(k+1)!} \frac{1}{m} + O(\frac{1}{m^2}) = \frac{1}{2(k-1)!} \text{ for } 1 \leq k < m.
\]
Taking the difference from the $k + 1$ summand of the Taylor series and letting $m \to \infty$ we obtain
\[
e - (1 + \frac{1}{m})^m = \frac{1}{2} \left( \sum_{k=1}^{m} \frac{1}{(k-1)!} \right) \frac{1}{m} + O(1/m^2) = \frac{e}{2m} + O(1/m^2)
\]
similarly
\[
e^{-1} - (1 - \frac{1}{m})^m = e^{-1} - \frac{1}{2m} + O(1/m^2)
\]
We conclude that
\[
[(1 + \frac{1}{m_R})^{m_R}(1 - \frac{1}{m_L})^{m_L}]^\gamma
\approx (e(1 - \frac{1}{2m_R}))^\gamma(e^{-1}(1 - \frac{1}{2m_L}))^\gamma
\approx [1 - \frac{c(\sqrt{L/R} + \sqrt{R/L})}{2\sqrt{h}}]^\gamma \sim e^{-c^2/2}
\]
The upshot of all these calculations is that
\[
B(h, Rh + c\sqrt{LRh}) = \sim \frac{1}{\sqrt{CRLh}}(\frac{L}{R})^\gamma e^{-c^2/2} 2^{H(L)h}
\]
Having obtained an estimate on the number of length $h$ sequences with $m = Rh + \gamma$ bits which are 1, we note that $B_{n,m+1}/B_{n,m} = \frac{n-m}{m+1} \sim \frac{1}{R}$ and similarly $B_{n,m+k}/B_{n,m} \sim (\frac{L}{R})^k$ and hence the number of addresses with at least $Rh + \gamma$ bits set to 1 is
\[
\sum_{\ell \geq Rh + \gamma} B_{n,\ell} \sim \frac{R}{R - L} \frac{1}{\sqrt{CRLh}}(\frac{L}{R})^\gamma e^{-c^2/2} 2^{H(L)h}
\]
The different address bits in the binary bias model are drawn independently. Each bit, say the $i$'th, is given by a random variable $X_i$ such that the probability that $X_i = 0$ is $L$ and that $X_i = 1$ is $R$. The probability of generating a given address with $Rh + \gamma$ bits set to 1 and the rest to 0 in the
bias model is given by $L^{Lh-\gamma} R^{Rh+\gamma}$ and so the total probability $Pr(Rh+\gamma_c)$ of all such addresses satisfies

$$Pr(Rh + c\sqrt{LRh}) \sim \frac{1}{\sqrt{CLRh}} e^{-c^2/2}$$

Given some $\varepsilon > 0$ we see that the probability $dP(c, \varepsilon, R)$ of having between $Rh + \gamma_c$ and $Rh + \gamma_{c+\varepsilon}$ bits set to 1 satisfies

$$dP(c, \varepsilon, R) \sim \varepsilon e^{-c^2/2}$$

Let $Y_h = \frac{\sum_{i=1}^h X_i - E(\sum_{i=1}^h X_i)}{\sqrt{\sigma^2(\sum_{i=1}^h X_i)}}$. We conclude from the computations above that $P_c = \lim_{h \to \infty} Pr(\{Y_h\} \leq c) = \frac{1}{\sqrt{C}} \int_{-\infty}^c e^{-x^2/2} dx$. The weak law of large numbers tells us that $\lim_{c \to \infty} P_c = 1$, hence,

$$\frac{1}{\sqrt{C}} \int_{-\infty}^\infty e^{-x^2/2} dx = 1$$

To calculate $C$, let $D = \int_{-\infty}^\infty e^{-x^2/2} dx$ so that $D^2 = C$. Consider the double integral $D^2 = \int_{-\infty}^\infty \int_{-\infty}^\infty e^{-x^2/2} e^{-y^2/2} dxdy$. We recall that in polar coordinates we have after computing the Jacobian of the coordinate transformation that $dxdy = rdrd\theta$ and so $C = D^2 = \int_0^{2\pi} \int_0^\infty re^{-r^2/2} drd\theta = 2\pi \int_0^\infty re^{-r^2/2} dr = 2\pi e^{-r^2/2}|_0^\infty = 2\pi$ giving us the full Stirling formula and also the full central limit theorem for binary random variables.

From the calculations above, we see that the hit ratio $Hit(Rh+\gamma_c)$ upon placing all addresses with at least $Rh + \gamma_c$ bits set to 1 is given by

$$Hit(Rh + \gamma_c) = \phi(c) = \frac{1}{C} \int_{-\infty}^c e^{-x^2/2} dx$$

as stated in the theorem. q.e.d.

## 1. Lorentzian geometry

We prove all the results which were left unproved in the text.

**Proof of lemma 9:** We first note that if $\gamma$ is a linear function of $s$ then the formula is exact. More generally, for a given $\varepsilon > 0$, consider events $A = E_0 \leq_M E_1 \leq_M \ldots \leq_M E_m = B$ on $\gamma$ such that $\ell_M(\gamma) + \varepsilon/2 > \sum_{j} d_M(E_i, E_{i+1})$. Assume $E_j = \gamma(s_j)$. Since the curve $\gamma$ is twice differentiable, each coordinate derivative $v_i'(s)$ is continuous and hence for any $\eta > 0$ we can find a division of the interval $s_j$, $s_{j+1}$ into sub-intervals with endpoints $s_j = s_{j,0} < s_{j,1} < \ldots < s_{j,k_j} = s_{j+1}$ such that on each sub-interval $I_{j,l} = [s_{j,l}, s_{j,l+1}]$ we have for all coordinates $i$, $Max_{I_{j,l}} v_i'(s) - Min_{I_{j,l}} v_i'(s) < \eta$. Let $E_{j,l} = \gamma(s_{j,l})$. According to the fundamental theorem of calculus we have $E_{j,l+1} - E_{j,l} = \int_{s_{j,l}}^{s_{j,l+1}} v(s) ds$. Let $v(j,l)$ be the vector $\frac{1}{s_{j,l+1} - s_{j,l}} \int_{s_{j,l}}^{s_{j,l+1}} v(s) ds$. By construction, the vector valued function $\tilde{v}$ which
takes the value \( v(j,l) \) on the interval \( (s_{j,l}, s_{j,l+1}] \) is the derivative of the piecewise linear curve \( \tilde{c} \) passing linearly between the points \( E_{j,l} \). Since these points contain the points \( E_j \) we have

\[
(211) \quad \ell_M(\gamma) \leq \ell_M(\tilde{c}) = \int_a^b \sqrt{\hat{v}(s)M\hat{v}(s)} \, ds < \ell_M(\gamma) + \varepsilon / 2
\]

On the other hand we have by the mean value theorem for integrals that for any coordinate \( i \) and any \( s, |v_i(s) - \tilde{v}_i| < \eta \). Since \( M \) (which is constant) and \( v_i(s) \) (the derivative of a Lifschitz function) are bounded we can find \( \eta \) small enough so that the condition \( |v_i(s) - \tilde{v}_i| < \eta \) will ensure that \(|v(s)M\hat{v}(s) - \tilde{v}M\tilde{v}| < \varepsilon / 2\) for all \( s \). This together with the estimate (211) proves the desired formula. \( q.e.d. \)

Assuming some basic results in the theory of Lebesgue integration it can be shown that formula (28) makes sense and gives the curve length of general causal curves without the differentiability assumptions. We outline the argument. We recall that a causal curve, parametrized by \( t \), the time coordinate of Minkowski space is Lifschitz.

Since \( f_n \) is continuous, we can find a piecewise constant function \( h_n \) with the same bound such that \( \int_a^b |f_n - h_n|^2 \, ds < 1/n \). We have the triangle inequality \( \sqrt{\int_a^b |g - f_n|^2 \, ds} + \sqrt{\int_a^b |f_n - h_n|^2 \, ds} \geq \sqrt{\int_a^b |g - h_n|^2 \, ds} \) from which we deduce that \( \sqrt{\int_a^b |g - h_n|^2 \, ds} \to 0 \).

Consider the function \( g \) restricted to some interval \( J = [\alpha, \beta] \). The constant \( h \) which minimizes \( \int_\alpha^\beta |g(s) - h|^2 \, ds \) will be the average value of \( g \) in \( J \), i.e., \( h = (\int_\alpha^\beta g(s) \, ds) / (\beta - \alpha) \). We can thus assume that in the functions \( h_n \), in any interval \( J \) in which \( h_n \) is constant it is given by the average value of \( g \) in \( J \).

For a finite interval \([a,b] \), and sequences of functions \( g_n, h_n \) all with a universal bound \( L \) and such that \( \int_a^b |g_n - f_n|^2 \, ds \to 0 \) we also have \( \int_a^b |g_n - f_n| \, ds \to 0 \) with the reverse implication also holding. In fact both statements are equivalent to the statement that for any \( \delta > 0 \), the measure of the set \( Y_{\delta,n} = \{ s \mid |g_n(s) - f_n(s)| \geq \delta \} \) tends to zero. This is because for all \( \delta > 0 \) we have the estimates \( \delta \mu(Y_{\delta,n}) \leq \int_a^b |g_n - f_n| \, ds \leq \delta (b - a) + (\mu(Y_{\delta,n}))2L \) and \( \delta^2 \mu(Y_{\delta,n}) \leq \int_a^b |g_n - f_n|^2 \, ds \leq \delta^2 (b - a) + (\mu(Y_{\delta,n}))(4L^2) \).

With these preliminaries we define \( s_j, E_j \) as in the previous proof. We take \( g_i = \gamma_i'(s) \) and we add the points \( s_{i,n,l} \) as endpoints of subintervals in which the approximating piecewise constant functions \( h_{i,n} \) are constant. By the Lifschitz condition all the \( g_i \) are bounded by \( c \). Let \( E_{i,n,l} = \gamma(s_{i,n,l}) \). By construction the points \( E_j \) are among the points \( E_{n,l} \) and by the fundamental theorem of calculus, the points \( E_{i,n,l} \) are on the curve \( \gamma \). The theorem now proceeds as before using the fact that the functions \( g_i \) and the entries of \( M \) are all bounded by some constant \( L \).

**Proof of theorem 15:**
To establish this property, we consider a sequence of events $E_i$ on the curve $\gamma$ such that $\sum_{i=1}^{n-1} d_M(E_i, E_{i+1}) < \ell(\gamma) + \varepsilon/2$. Such a sequence exists by the definition of the length. As usual we assume that the quadratic form $M$ is the standard one $t^2 - \sum_i x_i^2$. For a given $\delta > 0$ let $F_i = E_i - (\delta, 0, \ldots, 0)$ and $G_i = E_i + (\delta, 0, \ldots, 0)$. It is obvious that $F_i < M E_i < M E_{i+1} < M G_{i+1}$. By continuity of the quadratic form, if we choose $\delta$ small enough then we will have $d_M(F_i, G_{i+1}) < d_M(E_i, E_{i+1}) + 2\varepsilon/5$. We define the open set $U = U(\gamma, \varepsilon)$ to be the union of all the open intervals $I_i = I_M^0(F_i, G_{i+1})$.

Given a parametrized causal curve $\beta$ in $U$ it will meet some of the open sets $I_i$ in parameter intervals which together form an open covering of $[a, b]$, thus the length of $\beta$ will be at most the sum of lengths of the $\beta_i = \beta \cap I_i$. We have seen that for any causal curve $\beta_i$ in $I_i$ we have $\ell(\beta_i) \leq T(F_i, G_{i+1})$ and hence

$$\ell(\beta) \leq \sum_i \ell(\beta_i) \leq \sum_i d_M(F_i, G_{i+1}) \leq (\sum_i d_M(E_i, E_{i+1})) + \varepsilon/2 < \ell(\gamma) + \varepsilon$$

as required. q.e.d.

**Proof of lemma 10:**

Let $\bar{x} = (\bar{x}_0, \ldots, \bar{x}_n)$ be a point in a chart $\phi_i(U_i)$ of a space-time $S$ with Lorentz metric $g$ and assume that $g_{\bar{x}} = D_{a,1}$. As in our discussion of Minkowski space we will rename the first coordinate in this coordinate system $t$ and still use $x_1, \ldots, x_n$ for the other coordinates. Let $H$ be the hyperplane given by $t = 1$ with coordinates $x_1, \ldots, x_n$. For a given $r > 0$ consider the diagonal matrix $D_r$ with entries $(r, -1, \ldots, -1)$. Note that $g_{\bar{x}} = D_1$. It is clear that the causal cone corresponding to $D_r$ intersects $H$ in $B_r$, the ball of radius $r$ in $H$. Let $r_1 < 1 < r_2$. It is clear that for any $v \in H$ and in particular any $v \in B_{r_2}$ we have $vD_{r_1} v^t < vD_2 v^t < vD_{r_2} v^t$. The expression $vM v^t$ is continuous on the product space of vectors and matrices. Consequently, for each $v \in B_{r_2}$ we can find an open neighborhood $v \in U_v \subset V$ of vectors, and an open neighborhood $W_v$ of matrices $D_1 \subset W_v \subset P$ such that for any pair $(u, M) \in U_v \times W_v$ we have $uM v^t < uD_{r_2} v^t$. The sets $U_v$ form an open covering of the compact set $B_{r_2}$ hence there is a finite sub-cover $U_{v_1}, \ldots, U_{v_m}$. Taking $W_1 = \cap_{i=1}^m W_{v_i}$ we see that for any $v \in B_{r_2}$ and any $M \in W_1$ we have $vM v^t < vD_{r_2} v^t$. By the same argument we can find a neighborhood $W_2$ such that for all $M \in W_2$ and all $v \in B_{r_2}$, we have $vD_{r_2} v^t < vM v^t$. We define $W = W_1 \cap W_2$. Since the mapping $x \to g_x$ which defines a Lorentzian metric is continuous we see that we can find an open neighborhood $\bar{x} \in N_{\bar{x}}$ such that for all $x \in N_{\bar{x}}$ we have $g_x \in W$. Together with the defining properties of $W$ this implies that for all $x \in N_{\bar{x}}$ we have the following containments of the sets of causal vectors

$$(212) \quad C(D_{r_1}) \subset C(g_x) \subset C(D_{r_2})$$

In particular, this means that a differentiable parametrized causal curve $\gamma$ in $N_{\bar{x}}$ is also causal w.r.t. Minkowski space associated with the matrix $D_{r_2}$.
If we assume that $S$ is strongly causal, we may, by further restriction, assume that the neighborhood $N_{\delta}$ is causally convex.

We assume that we are given a Euclidean metric associated with a matrix $K$ on $\phi_i(U_i)$.

Fix some $\varepsilon > 0$. Consider the vector $\tau = (1, 0, \ldots, 0)$. Choose some $\rho < 1$ such that $\sqrt{1 - \rho^2} < d_K(\tau)\varepsilon / 2$. Obviously the last inequality implies that $\sqrt{1 - \rho^2} < d_K(\tau)\varepsilon / 2$ for all $1 \geq r \geq \rho$. Now choose $r_1 < 1 < r_2$, with $r_1, r_2$ close enough to 1 so that the inequality $\sqrt{r_2^2 - \rho^2} < d_K(\tau)\varepsilon$ holds and such that $\sqrt{r_1^2 - \rho^2} > 1 - \varepsilon$. Construct the neighborhood $N_{\varphi,\varepsilon}$ as above using the ($\varepsilon$ dependent) chosen $r_1, r_2$. Consider a vector $v \in H$ given by $v = (1, x_1, \ldots, x_n)$ and let $\tau = \sqrt{x_1^2 + \ldots + x_n^2}$. Since the expression $\sqrt{r_2^2 - r^2}$ decreases with $r$ we have for all $r \geq \rho$ and any $y \in N_{\varphi,\varepsilon}$ that

$$\frac{\sqrt{v g_y v^t}}{d_K(\tau)} \leq \frac{\sqrt{r_2^2 - r^2}}{d_K(\tau)} < \varepsilon$$

Since the ratio $\sqrt{r_2^2 - r^2}$ is also a decreasing function of $r$ we conclude that with these choices, for all $0 \leq r \leq \rho$, we have $\sqrt{r_2^2 - r^2} > 1 - \varepsilon$. Let $v = (t, x_1, \ldots, x_n)$ be a non-zero causal vector w.r.t. $D_{r_2}$, in particular $t \neq 0$. Let $\tau_v = (t, 0, \ldots, 0)$ and let $y \in N_{\varphi,\varepsilon}$. The above inequalities imply that either

$$(213) \quad \frac{\sqrt{v g_y v^t}}{d_K(\tau_v)} < \varepsilon$$

when $r \geq \rho$, or

$$(214) \quad \sqrt{v D_{r_2} v^t} \geq \sqrt{v g_y v^t} \geq \sqrt{v D_{t} v^t} \geq \sqrt{v D_{r_2} v^t} \frac{1}{1 - \varepsilon}$$

when $r < \rho$. Stated otherwise, tangent vectors of causal curves in $N_{\varphi,\varepsilon}$ are either very short according to the Lorentz metric (almost light-like) or their length is very well approximated (multiplicatively) by the approximating Minkowski metrics. Equation (214) completes the proof of the lemma. \textit{q.e.d.}

**Proof of theorem**: 16:

We begin by showing that $C_D(A, B)$ is sequentially compact in the Hausdorff metric. Each point in $D$ is contained in a normal open set $V_x$. Since the $V_x$ form a covering of $D$ and $D$ is compact there is a finite collection $V_1 = V_{x_1}, \ldots, V_m = V_{x_m}$ which covers $D$. Let $\gamma : [a, b] \rightarrow D$ be a parametrization of a causal curve in $D$. Let $U_{i,\gamma} = \gamma^{-1}(V_i)$. The $U_{i,\gamma}$ form a cover of $[a, b]$. We construct a minimal sub-cover $W_{j,\gamma}$, $1 \leq j \leq m(\gamma)$, of $\gamma$ by sets of the form $U_{i,\gamma}$ as follows. We consider all sets $U_{i,\gamma}$ which contain the lower endpoint $a$ of $I$. Among these, we choose the interval $W_{1,\gamma} = U_{i_1,\gamma} = [a, b_{i_1}]$ which has the largest upper endpoint $b_{i_1}$. If there are ties we break them arbitrarily. We remove all other intervals $a \in U_{i,\gamma}$ from the list, since they
are not needed for covering \( I \). We consider the set \( A_1 \) consisting of all remaining open intervals such that \( b_{i_1} \in U_{i, \gamma} \). Since the \( U_{i, \gamma} \) formed a covering and all the intervals we have removed do not contain \( b_{i_1} \) by construction, \( A_1 \) is not empty. We choose from \( A_1 \) the interval \( W_{2, \gamma} = U_{i_2, \gamma} \), with the largest upper endpoint \( b_{i_2} \). We then repeat the process inductively until we obtain the desired covering. It is clear from the construction that \( W_{i, \gamma} \) intersects \( W_{j+1, \gamma} \) and that for any \( j \leq l - 2 \) and any \( x \in W_{j, \gamma} \) and \( y \in W_{l, \gamma} \) we have \( x < y \). We call \( \gamma(b_{i_j}) \) the exit point of \( \gamma \) from \( W_{j, \gamma} \). Consider the mapping \( F \) which sends a curve \( \gamma \) to the sequence \( i_1, ..., i_m(\gamma) \) such that \( W_{j, \gamma} = U_{ij, \gamma} \).

Let \( \gamma_n \in C_D(A, B) \) be a sequence of causal curves. We want to show that it contains a convergent sub-sequence. Since the image of \( F \) is finite there is a subsequence \( \gamma_{n_k} \) on which \( F \) is constant. By passing to such a subsequence we may assume, after some re-indexing and without loss of generality that each \( \gamma_n \) is minimally covered as above by \( W_{1, \gamma}, ..., W_{m, \gamma} \) were \( W_{j, \gamma} = \gamma^{-1}(V_j) \). Let \( E_{0,n} = \gamma_n(a_n) \) and \( E_{m,n} = \gamma_n(b_n) \). In addition, for \( j = 1, ..., m - 1 \), let \( E_{j,n} = \gamma_n(s_{j,n}) \), be the exit point of \( \gamma_n \) from \( V_j \), which means that \( W_{j,n} \) has the form \((w_{j,n}, s_{j,n})\) for some \( w_{j,n} \). Since \( D \) is compact we can find a sub-sequence \( \gamma_{n_k} \) of \( \gamma_n \) such that for any \( 0 \leq j \leq n \), the subsequence \( s_{j,n_k} \) converges to \( s_j \). From now on we assume that we have restricted ourselves to such a sub-sequence.

We want to consider the curves \( \gamma_n \) between \( E_{j-1,n} \) and \( E_{j,n} \), i.e., on the interval \([s_{j-1,n}, s_{j,n}]\) as curves in the closure \( \tilde{V}_{x_j} \) of \( V_{x_j} \) with a parametrization similar to the local time parameter \( t_j \) corresponding to the majorizing Minkowski metric \( M_{majo,j} \) of the normal neighborhood. The motivation is to consider the \( \gamma_n \) as Lifschitz functions. However, we would also like to re-parametrize the \( \gamma_n \) so that all \( \gamma_n \) restricted to the interval \([s_{j-1,n}, s_{j,n}]\) will have a common parametrization interval \( I_j \) so we can apply the Arzela-Ascoli theorem, which is stated in the appendix. Both goals will be achieved by linearly stretching the local time parameter by different amounts for each pair \( j,n \) without destroying equi-continuity.

We will re-parametrize each \( \gamma_n \). We will define inductively \( \tau_j \), independent of \( n \) such that after re-parametrization we will have \( \gamma_n(\tau_j) = E_{j,n} \) for any \( n \). In addition, after re-parametrization \( \gamma_n \) restricted to the interval \([\tau_{j-1}, \tau_j]\) will be Lifschitz with a constant \( c_j \) which is independent of \( n \). We let \( \tau_0 = 0 \) and define the other values \( \tau_j \) inductively.

Given some \( j = 1, ..., m \), we consider two possible cases. In the first case which we call the non-degenerate case we have \( E_{j-1} \neq E_j \). For a point \( E \in \tilde{V}_{x_j} \), let \( t_j(E) \) denote the value of the \( t_j \) coordinate at \( E \). Since \( \gamma_n \) is a causal curve, w.r.t. \( M_{majo,j} \), by the properties of the \( t_j \) coordinate on \( \tilde{V}_j \) we know that for points \( E \in \gamma_n \) such that \( E_{j-1} \leq E \leq E_j \) the map \( E \rightarrow t_j(E) \) is 1-1.

We define \( \tilde{s}_n(E) = \tilde{t}_{j,n}(E)(t_j(E_j) - t_j(E_{j-1})) \) where

\[
\tilde{t}_{j,n}(E) = \frac{t_j(E) - t_j(E_{j-1,n})}{t_j(E_{j,n}) - t_j(E_{j-1,n})}
\]
As $E$ varies between $E_{n,j-1}$ and $E_{n,j}$, $\ell_{j,n}(E)$ varies continuously and strictly monotonically from 0 to 1 and hence $\delta_n$ strictly increases continuously from 0 to $t_j(E_j) - t_j(E_{j-1})$. We define $\tau_j = \tau_{j-1} + t_j(E_j) - t_j(E_{j-1})$ and define the new parametrization to be such that

$$\gamma_n(\tau_{j-1} + \delta_n(E)) = E$$

(215)

We recall that causal curves in $\bar{V}_j$ parametrized by $t_j$ are Lifschitz with constant $c_j$. The new parameter for $\gamma_n$ in the interval $\tau_{j-1}, \tau_j$ has the form $d_{n,j} + e_{n,j}t_j$ where $e_{n,j} = \frac{t_j(E_j) - t_j(E_{j-1})}{t_j(E_{n,j}) - t_j(E_{j-1,n})}$. Shifting the parameter does not change the Lifschitz constant, while scaling it can change the constant from $c_j$ to $c_j/e_{n,j}$. Thus, we will have a Lifschitz constant independent of $n$ if $1/e_{n,j} = \frac{t_j(E_{n,j}) - t_j(E_{j-1,n})}{t_j(E_{n,j}) - t_j(E_{j-1,n})}$ is bounded. However, $t_j$ is bounded on all of $\bar{V}_j$ so we get a bound that is independent of $n$.

In the second case we will call the degenerate case, we have $E_{j-1} = E_j$. If $E_{j-1,n} \neq E_{j,n}$ we define $\delta_n(E) = \frac{t_j(E) - t_j(E_{j-1,n})}{t_j(E_{n,j}) - t_j(E_{j-1,n})}$ as $E$ varies from $E_{j-1,n}$ to $E_{j,n}$ the value of $\delta_n$ increases continuously from 0 to 1. We define $\tau_j = \tau_{j-1} + 1$ and use equation (215) to parametrize the curve. Again, the new parameter has the form $d_{n,j} + e_{n,j}t_j$ with

$$\frac{1}{e_{n,j}} = t_j(E_{n,j}) - t_j(E_{j-1,n}) \to 0$$

hence the Lifschitz condition still holds independently of $n$.

For $0 < j < m$, the point $E_{j-1,n}$ lies in the interior of $\bar{V}_j$ while the point $E_{j,n}$ is on the boundary of $\bar{V}_j$, hence they cannot be the same and the above parameter construction is well defined. On the other hand it may happen in a very degenerate situation that $E_{m-1,n} = E_{m,n}$. In this case we again define $\tau_m = \tau_{m-1} + 1$ and map all the interval $[\tau_{m-1}, \tau_m]$ to the point $E_{m-1,n} = E_{m,n}$. Since a constant function is obviously Lifschitz with any Lifschitz constant, say 1, we still have a universal Lifschitz constant good for all $n,j$.

Armed with the re-parameterizations we let $\gamma_{j,n}$ denote $\gamma_n$ restricted to $I_j = [\tau_{j-1}, \tau_j]$ as a function to the compact space $\bar{V}_j$, $j = 1, ..., m$. Since these functions are Lifschitz functions with a constant independent of $n$ we can repeatedly apply the Arzela-Ascoli theorem to each of the sequences $\gamma_{j,n}$ to find a sub-sequence $\gamma_{n_k}$ which for each interval $I_j$ converges uniformly with respect to $d_j$, the Euclidean metric on $\bar{V}_j$ to a curve $\gamma_j$. We claim that the $\gamma_{n_k}$ will converge to $\gamma$, the concatenation of the $\gamma_j$, uniformly w.r.t. $d_S$ on the entire parameter interval $[\tau_0, \tau_m]$. Consider some $\varepsilon > 0$ and some $j$. Assume that for each $\delta > 0$ we can find some point $E_\delta \in \gamma_j$ such that the ball $B_{Euc,j}(E_\delta, \delta)$ is not contained in the ball $B_S(E_\delta, \varepsilon)$. Taking a sequence $\delta_h \to 0$, the sequence of points $E_{\delta_h}$ will have a subsequence converging to $E \in \bar{V}_j$. It is easy to see that for $E$, no ball $B_{Euc,j}(E, \eta)$ is contained in $B_S(E, \varepsilon)$, contradicting the equivalence of the topologies $d_S$ and $d_{Euc,j}$ in a chart containing a neighborhood of $\bar{V}_j$. We conclude that for some $\delta_j > 0$
and any $E \in \gamma_j$ we have $B_{\text{Haus}}(E, \delta_j) \subset B_S(E, \varepsilon)$. Take $\delta_\varepsilon = \text{Min}_j \delta_j$ and let $k_0$ be such that for all $k > k_0$ and all $s \in [\tau_0, \tau_m]$, $d_{\text{Haus}}(\gamma(s), \gamma_{n_k}(s)) < \delta$. Then for all $s \in [\tau_0, \tau_m]$ we have $d_S(\gamma(s), \gamma_{n_k}(s)) < \delta$ and hence $\gamma_{n_k} \to \gamma$ in the Hausdorff metric of $S$, proving sequential compactness and compactness of $C_D(A, B)$ in this metric.

Consider an open covering of $C_D(A, B)$ in the $C_0$ topology. Let $N$ be an open set in $S$ and $\gamma$ a causal curve such that $\gamma \subset N$. We claim that there is a $\varepsilon(N) > 0$ such that $B_S(\gamma, \varepsilon(N)) \subset N$. If not, there is a sequence $x_n \in S$ and points $E_n \in \gamma$ such that $d_S(x_n, E_n) < 1/n$ and $x_n \notin B_S(\gamma, 1/n)$. Considering a subsequence $E_{n_k}$ which converges to some point $E \in \gamma$ we see that no open ball around $E$ can be contained in $N$, a contradiction to $N$ being open. We conclude that $B_{\text{Haus}, S}(\gamma, \varepsilon(N)) \subset C_N(A, B)$. Taking the open ball $B_{\text{Haus}, S}(\gamma, \varepsilon(N))$ in the Hausdorff metric for all pairs $\gamma, N$ such that $C_N(A, B)$ participates in the open covering, we obtain an open covering in the Hausdorff metric. Since $C_D(A, B)$ is compact in the Hausdorff metric, there is a finite sub-cover by $B_{\text{Haus}, S}(\gamma_i, \varepsilon(N_i))$, $i = 1, \ldots, r$. The corresponding sets $C_{N_i}(A, B)$ will form a finite sub-cover proving the compactness of $C_D(A, B)$ in the $C_0$ topology. q.e.d.

**Proof of theorem 26:**

Instead of sampling $n$ points (passengers) we will assume a sampling of $P(n)$ passengers. Using de-Poissonization arguments as before we can see that it is enough to establish that w.h.p. $|\hat{L}_{n,h,l,w,p} - \hat{L}_{n,k,p}| \leq \varepsilon \sqrt{n}$.

Consider two passengers represented by coordinates $X = (q, r)$ and $X' = (q + dq, r - dr)$, $dq, dr > 0$. Consider the time when passenger $X$ arrives at his/her designated row. Consider the $f$ passengers $X_1, \ldots, X_f$ which are standing in the aisle behind $X$ but in front of $X'$. Assume that $X$ blocks $X'$. According to (24) this means that $fw/l > row_X - row_{X'}$, where $row(X)$ is the actual row in the airplane assigned to passenger $X$. Let $X_j$ be the first passenger behind $X$ whose coordinate satisfies $r_{X_j} \leq r - dr$. Since $X_j$ cannot reach his/her row we conclude that $fw/l \geq (j - 1)w/l \geq row_X - row_{X_j} \geq row_X - row_{X'}$. By definition, the passengers $X_1, \ldots, X_{j-1}$ have coordinates which satisfy $r_{X_i} \geq r - dr$ and $q \leq q_{X_i} \leq q + dq$, $i = 1, \ldots, j - 1$. According to Poisson sampling the number of passengers which satisfy these coordinate constraints has distribution $P(Cn)$ where

$$C = C(q, r, dq, dr) = \int_{r-dr}^1 \int_{q+dq}^q p(u, z)du dz$$

Recall that each such passenger occupies $w$ aisle length, or $w/l$ of the distance between successive rows. On the other hand, $row_X - row_{X'}$ is the number of passengers $Z$ with row coordinate satisfying $r - dr < r_Z < r$, divided by the number of passengers per row which is $h$. Since, we have normalized the distribution so that $\int_0^1 p(q, r) dq = 1$, the distribution of the number of such passengers is $P((dr)n)$. We would like to consider a closely related random variable that is independent from the previous one.
and so we consider the number of passengers $Y$ whose $q$ coordinate satisfies either $q_Y \leq q$ or $q_Y \geq q + dq$. This random variable has distribution $P((dr - \int_{r-dr}^r \int_q^{q+dq} p(u,v) dudv)n)$. Let

$$E = E(q,r,dq,dr) = dr - \int_{r-dr}^r \int_q^{q+dq} p(u,v) dudv$$

We conclude from our discussion that the probability that $X$ will block $X'$ is at most the probability that

$$P(\frac{\int_1^r \int_q^{q+dq} p(u,z) dudz n}{w/l} > P((dr - \int_{r-dr}^r \int_q^{q+dq} p(u,v) dudv)n)/h)$$

or equivalently

$$kP(Cn) > P(En)$$

where the distributions are sampled independently.

Let $X, X'$ be a pair of points and assume that $Max(Cn, Dn) > n^\nu$ for some $\nu > 0$ and $n$ large enough. Assume further that

$$\beta kCn < En$$

for some $\beta > 1$. Then, by (40), the probability that $X'$ blocks $X$ will be at most $e^{-n\nu}$ for some $\mu > 0$. Consequently, if the number of possible pairs $X, X'$ will be polynomial in $n$ then w.h.p. non of the pairs will be blocking.

At the point $(q,r)$ we have that a vector $v = (v_1, v_2)$, with $v_1 > 0$ is causal w.r.t. the boarding metric if and only if $k\alpha(q,r)v_1 \geq -v_2v_1$, or equivalently, $k\alpha(q,r)v_1 \geq -v_2$. Applying to $v = (dq, -dr)$ we obtain the condition $k\alpha(q,r)dq \geq dr$.

Consider a covering of $S_{k,p}$ by $\delta$-normal neighborhoods $N_j$, for some $\delta > 0$. As in the proof of theorem 19, we can define a partial order $\leq_{\text{max}}$ assigned to the cover, where $X \leq_{\text{max}} Y$ if and only if there is a sequence of points $X = X_1, X_2, ..., X_k = Y$ such that any successive pair $X_k, X_{k+1}$ belong to some neighborhood $N_j(k)$ and satisfy $X_k \leq_{\text{max},j} X_{k+1}$. The proof of theorem 19 shows that we can choose the covering so that the longest chain w.r.t. $\leq_{\text{max}}$ has w.h.p. length at most $(\text{diam}(S_{k,p}) + \varepsilon/2)\sqrt{n}$. We assume that we have chosen a covering that satisfies this property.

We show that w.h.p., there are relatively few pairs $X, X'$, such that, $X$ blocks $X'$, but $X \leq_{\text{max}} X'$ is not satisfied. We will call such pairs, renegade pairs. We also claim that w.h.p., any such pair has unusually small values of $dq, dr$, hence, they are very close to each other.

The point $X$ is in some neighborhood $N_j$. We first assume that $X'$ is there as well. By the construction of majorizing Minkowski metrics, the condition that, $X \leq_{\text{max},j} X'$ does not hold, will be more stringent than in the boarding metric and hence will take the form $(k\alpha(q,r) + \omega_j(q,r))dq \leq dr$, where $\omega_j(q,r) > 0$ is continuous in $(q,r)$, at points where the maximizing metric is defined. We let $\omega > 0$ be the minimum of $\omega_j(q,r)$ on $\bar{N}_j$, the
closure of $N_j$. Let $\omega = \text{Min}_j \omega_j/2$. For this choice we have that $X \leq \max X'$ does not hold only if
\begin{equation}
(219) \quad dr \geq k\alpha(q, r) + \omega
\end{equation}
hence, a renegade pair must satisfy this condition.

For a given $x > 0$, let
\[ f(x) = \text{Max}_{q,r,y} |p(q + y, r) - p(q, r)| \]
with $q, r$ arbitrary and $y < x$. By definition $f(x)$ is a non-decreasing function, and since $p$ is (uniformly) continuous, we have $\lim_{x \to 0} f(x) = 0$. In terms of $f$ we have the following estimate
\[ \int_r^1 \int_q^{q + dq} p(u, v) dudv = \int_r^1 \int_q^{q + dq} (p(u, v) - p(u, q)) + p(u, q) dudv \]
\[ \leq (1 - r) f(dq) dq + \alpha(q, r) dq \leq (f(dq) + \alpha(q, r)) dq \]
Let
\[ \gamma = \text{Max}_{q,r} p(q, r) \]
We choose an $\eta > 0$ small enough so that
\[ \frac{\omega}{2} > k f(\eta) + 2(k + 1)^2 \gamma^2 \eta \]

We will consider the existence of renegade pairs in several cases.

**Case I:** Assume $X, X'$ are such that
\begin{equation}
(220) \quad n^{-99/100} < dq < \eta
\end{equation}
In particular $dq n \geq n^{1/100}$. We consider two sub-cases.

We first assume that
\begin{equation}
(221) \quad dr \geq (2(k + 1)\gamma) dq
\end{equation}
Then,
\[ En = dr - \int_{r-dr}^r \int_q^{q + dq} p(u, v) dudv \geq 2\gamma(k + 1)dqn - \gamma dq n = (2k + 1)\gamma dq n \]
On the other hand
\[ kCn = k \int_{r-dr}^r \int_q^{q + dq} p(u, z) dudzn \leq k\gamma dq n \leq \frac{k}{2k + 1} En \]
By (217,218), with $\beta = 2$, the probability that $X, X'$ which satisfy (220) and (221) is a renegade pair is at most $e^{-n\mu}$ for some $\mu > 0$. Since there are at most $n^2$ pairs of passengers we see that w.h.p., there are no renegade pairs of passengers with $dq, dr$ satisfying the assumptions. Moreover we claim that there is no renegade pair $X, X'$ satisfying the assumptions, regardless of whether $X, X'$ represent the coordinates of passengers. To see this, we round the coordinates of the points $X$ and $X'$ to produce points $Y, Y'$ whose coordinates are integer multiples of $n^{-995/1000}$. Denoting by
$d\bar{q}, d\bar{r}$ the coordinate differences between $Y$ and $Y'$ we have by definition $|dq - d\bar{q}|, |dr - d\bar{r}| \leq n^{-995/1000}$. For $n$ large enough, this implies

\begin{equation}
\frac{1}{2}n^{-99/100} < n^{-99/100} - 2n^{-995/1000} \leq dq
\end{equation}

and

\begin{equation}
d\bar{r} \geq (2k + 3/2)d\bar{q}
\end{equation}

Let $\tilde{C}, \tilde{E}$ be the integrals corresponding to the pair $Y, Y'$. The symmetric differences between the domains of integration which appear in the definitions of $C, E$ and those in $\tilde{C}, \tilde{E}$ respectively have area at most $cn^{-995/1000}$, for some constant $c$. Since the integrand is bounded by $\gamma$ we have $|E - \tilde{E}|, |C - \tilde{C}| < c\gamma n^{-995/1000}$.

If we have a renegade pair $X, X'$ then $kP(Cn) > P(En)$ and this implies w.h.p. $kP(\tilde{C}n) > 0.99P(\tilde{E}n)$. Repeating the arguments we had before the probability of such an event, given (222,223) is at most $e^{-n\tilde{\mu}}$ for some $\tilde{\mu} > 0$. Since there are at most $n^3$ pairs $Y, Y'$, with coordinates which are integer multiples of $n^{-995/1000}$, we see that, w.h.p. there is no pair with $kP(\tilde{C}n) > 0.99P(\tilde{E}n)$ and consequently, w.h.p. there is no renegade pair $X, X'$ as desired. In fact, the proof gives an exponentially small probability for such a pair.

In the second sub-case we assume that $dr \leq (2(k + 1)\gamma)dq$. Since we assumed that $dq < \eta$ we have $dr \leq (2(k + 1)\gamma)\eta$. We also have

\begin{align*}
kC &= k \int_{r-dr}^{1} \int_{q}^{q+dq} p(u, z)dudz \\
&= k \int_{r-dr}^{r} \int_{q}^{q+dq} p(u, z)dudz + k \int_{r-dr}^{1} \int_{q}^{q+dq} p(u, z)dudz \\
&\leq k\gamma dr dq + k(\alpha(q, r) + f(dq))dq \leq 2k(k + 1)\gamma^2 \eta dq + (k\alpha(q, r) + kf(\eta))dq
\end{align*}

On the other hand, using (219)

\begin{align*}
E &= dr - \int_{r-dr}^{r} \int_{q}^{q+dq} p(u, v)dudv \\
&\geq (k\alpha(q, r) + \omega)dq - \gamma dr dq \geq (k\alpha(q, r) + \omega)dq - (2(k + 1)\gamma^2 \eta)dq
\end{align*}

By the definition of $\eta$ and the above estimates, we have

\begin{align*}
En - kCn \geq (\omega - 2(k + 1)^2 \gamma^2 \eta - kf(\eta))dq n \geq \frac{\omega}{2} dq n
\end{align*}

We conclude as before, that w.h.p., there are no renegade pairs of points $X, X'$ with $dq, dr$ satisfying the assumptions.

**Case II:** $dq > \eta$. By the definition of $X$ blocking $X'$, there are passengers in the queue $X_1, ..., X_q$, between $X$ and $X'$ when $X$ reaches his/her seat. Let $\rho_i$ be the row location of $X_i$ while standing behind $X$ and $X_1, ..., X_{i-1}$ and let $\bar{X}_i$ be the point with coordinates $(q_i, \rho_i/m)$. We have that $\bar{X}_i$ blocks
\(X_j\), if \(j > i\). Consider a sequence \(X_{i_1}, X_{i_2}, \ldots, X_{i_k}\) such that for each consecutive pair \(X_{i,j}, X_{i,j+1}\) we have \(n^{-99/100} \leq dq \leq \eta \) and such that \(X_{i,j}, X_{i,j+1}\) belong to some mutual neighborhood \(N_j\). We round the coordinates of the points \(X_i\) to produce points \(Y_i\) whose coordinates are multiples of \(n^{-995/1000}\). Consider a potential renegade pair \(X, X'\). Since \(X \leq_{\max} X'\) does not hold there is some pair \(X_{i,j}, X_{i,j+1}\) for which \(X_{i,j} \leq_{\max} X_{i,j+1}\) does not hold and will produce a renegade pair. Since \(dq \geq n^{-99/100}\), the Euclidean distance between \(X_j\) and \(X_{j+1}\) has a larger order of magnitude than the distance between \(X_i\) and \(Y_i\). From this observation it is easy to verify that if \(X_{i,j}, X_{i,j+1}\) is a renegade pair and \(n\) is large enough, then for the pair \(Y_i, Y_{i,j+1}\), condition (219) holds since \(\omega = \min_j \omega_j / 2 < \min_j \omega_j\). Arguing as in the previous case we see that the probability of having a pair \(Y_{i,j}, Y_{i,j+1}\) satisfying (219) and being in such close proximity to a renegade pair is at most \(e^{-n^\mu}\) for some \(\mu > 0\). Since there are at most \(n^4\) pairs of \(Y\)'s with coordinates which are multiples of \(n^{-995/1000}\), then no such pair exists and hence no pair \(X, X'\) exists.

We consider the last case, \(dq < n^{-99/100}\). We consider two sub-cases. In the first sub-case we assume that \(dr \geq n^{-98/100}\). The latter implies that

\[
E = \left( \int r_{-dr}^r \int q + dq \ p(u, v) du dv \right) \geq \left( n^{-98/100} - \gamma n^{-99/100} \right) \geq \frac{1}{2} n^{-98/100}
\]

for \(n\) large enough. On the other hand

\[
kC = k \left( \int q + dq \ p(u, z) du dz \right) n \leq k\gamma dq \leq k\gamma n^{-99/100}
\]

By our previous arguments the probability that \(kP(Cn) > P(En)\) is at most \(e^{-n^\mu}\) for some \(c, \mu > 0\), hence, for \(n\) large enough, w.h.p., we will not have any pairs satisfying these conditions.

**Case III:** \(dr < n^{-98/100}\). We claim that w.h.p., there are fewer than \(12n^{3/100}\) pairs of passengers \(X, X'\) with \(dq \leq n^{-99/100}\) and \(dr \leq n^{-98/100}\).

To see this consider the rectangles

\[
R_{i,j} = \left[ \frac{2i}{n^{99/100}}, \frac{2i + 2}{n^{99/100}} \right] \times \left[ \frac{2j}{n^{98/100}}, \frac{2j + 2}{n^{98/100}} \right]
\]

for \(1 \leq i \leq n^{99/100}/2\) and \(1 \leq i \leq n^{98/100}/2\). Each pair of points \(X, X'\) with \(dq, dr\) as above belongs either to a rectangle of the form \(R_{i,j}\), or to a rectangle of the form \(R_{i,j} + \left( \frac{1}{n^{99/100}}, 0 \right)\), or to a rectangle of the form \(R_{i,j} + \left( 0, \frac{1}{n^{98/100}} \right)\) or to one of the form \(R_{i,j} + \left( 0, \frac{1}{n^{98/100}} \right)\). Since the area of a rectangle is \(4n^{-197/100}\), the number of points in a rectangle is distributed like \(P(4n^{-97/100})\). For this distribution the probability of having three points or more is approximately \(64n^{-291/100}/6\). Since there are only \(n^{197/100}/4\) rectangles \(R_{i,j}\), w.h.p., no rectangle contains 3 or more points. The probability of having 2 points is \(8n^{-194/100}\). Consider the binary variable
Consider a chain $C$ of passengers $X_1, ..., X_m$ with respect to the boarding relation. We construct a sub-chain $\tilde{C}$ w.r.t. $\leq_{\text{max}}$ as follows. We let $X_1 \in \tilde{C}$. Assume inductively that $X_i$ is the last element in $\tilde{C}$. Let $j > i$ be the smallest index for which $X_i \leq_{\text{max}} X_j$ and add $X_j$ to $\tilde{C}$. Obviously, $j \neq i + 1$ if and only if the pair $X_i, X_{i+1}$ is a renegade pair. Let $I$ be the set of indices $i$ such that $X_i, X_{i+1}$ is a renegade pair. As we have seen, w.h.p., $|I| \leq 12n^{3/100}$. Given such a renegade pair with $X_i = (q, r)$, The set of points $X$ such that $X \leq_{\text{max}} X$ has the form $r \geq G(\bar{q})$ for some function non-increasing function $G$ defined on the interval $[q, 1]$. Consider the function $G_s(\bar{q}) = G(\bar{q}) - n^{-98/100}$.

Let $D_i$ be the domain bounded between the graphs of $G_s$ and $G$ and let $A_i$ be the set of passengers in $D_i$. Since the Euclidean area bounded between these graphs is at most $n^{-98/100}$, the probability that $A_i \geq 2\gamma n^{2/100}$ is at most $e^{-n^\mu}$ for some $\mu > 0$. We conclude that, w.h.p. the union of all the sets $A_i$ has size at most $2\gamma n^{5/100}$.

We claim that w.h.p., for any $i \in I$, the points $X_{i+1}, ..., X_{j-1}$ all belong to $A_i$. To see why, consider the relation $\leq_{\text{max},s}$ which is defined by $X \leq_{\text{max},s} X'$ if and only if $X + (0, n^{-98/100}) \leq_{\text{max}} X' + (0, n^{-98/100})$. If $n$ is large enough then by the uniform continuity of $\alpha(q, r)$, this relation will still majorize the boarding relation, but with the constant $\omega$ replaced by a smaller constant say $\omega/2$. We can define renegade pairs with this relation replacing $\leq_{\text{max}}$. The probability that a point $X$ will participate in a renegade pair (in the new definition) is at most $12n^{-97/100}$. In addition, the definition of $D_i$ is independent of the events which determine renegade pairs w.r.t. $\leq_{\text{max},s}$, with the exception of the pair $X_i, X_{i+1}$ itself. We conclude that w.h.p., non of the points in any of the $A_i$ (except the $X_i$ themselves), will be part of a renegade pair w.r.t. $\leq_{\text{max},s}$ and hence, in any domain $D_i$, $X$ blocking $X'$ implies $X \leq_{\text{max},s} X'$. We have seen that w.h.p., all points $X_{i+1}$ of the renegade pairs satisfy $dr \leq n^{-98/100}$, which means that $X_{i+1} \in D_i$. Since $X_{i+1}$ blocks $X_{i+2}$ the latter will also be in $D_i$ since otherwise they will be a renegade pair w.r.t. $\leq_{\text{max},s}$ with $X_{i+1} \in D_i$. The argument continues for all elements up to $X_{j-1}$.

We have seen that w.h.p., $C$ is the union of a chain w.r.t. $\leq_{\text{max}}$ and a chain of size at most $2\gamma n^{-5/100}$ and thus we obtain the desired upper bound on the size of a maximal chain. The lower bound is proved in precisely the same way using the minorizing Minkowski metrics. $q.e.d.$

**Proof of theorem 44:**

We start with the following lemma.
Lemma 28. Consider the power series

\[ u_n(q) = \frac{q^n}{1-q^n} = q^n + q^{2n} + q^{3n} + \ldots \]

The function \( u_n = u_n(q) \) satisfies the following identities:

1) \( \sum_{m \geq 1} u_m(1 + u_m) = \sum_{n \geq 1} nu_n \)
2) \( \sum_{m \geq 1} (-1)^{m-1} u_{2m}(1 + u_{2m}) = \sum_{n \geq 1} (2n - 1)u_{4n-2} \)
3) \( u_l u_{k-l} = u_k(1 + u_l + u_{k-l}) \)
4) \( u_{k+l} + u_l u_{k+l} = u_k(u_l - u_{k+l}) \)

**Proof:** By definition \( u_m = q^m + q^{2m} + \ldots \), hence we have

\[ 1 + u_m = 1 + q^m + q^{2m} + \ldots = \frac{1}{1 - q^m} \]

Starting with the geometric series \( \frac{1}{1-q} = \sum_{n \geq 1} q^n \) and differentiating we get \( \frac{1}{(1-q)^2} = \sum_{n \geq 1} nq^{n-1} \) and multiplying by \( q \) we obtain \( \frac{q}{(1-q)^2} = \sum_{n \geq 1} nq^n \). Applying to \( q^m \) instead of \( q \) and summing over all \( m \) we obtain

\[
\sum_{m \geq 1} u_m(1 + u_m) = \sum_{m \geq 1} \frac{q^m}{(1 - q^m)^2} = \sum_{m \geq 1} \sum_{n \geq 1} nq^{nm} = \sum_{n \geq 1} nu_n
\]

To prove the second identity we compute

\[
\sum_{m \geq 1} (-1)^{m-1} u_{2m}(1 + u_{2m}) = \sum_{m \geq 1} \frac{(-1)^{m-1} q^{2m}}{(1 - q^{2m})^2} = \sum_{m \geq 1} \sum_{k \geq 1} (-1)^{m-1}kq^{2mk}
\]

\[
= \sum_{k \geq 1} \frac{kq^{2k}}{1 + q^{2k}}
\]

We further note that \( \frac{kq^{2k}}{1+q^{2k}} = \frac{kq^{2k}}{1-q^{2k}} - \frac{(2k)q^{2(2k)}}{1-q^{2(2k)}} \), hence

\[
\sum_{k \geq 1} \frac{kq^{2k}}{1 + q^{2k}} = \sum_{k \geq 1} \frac{kq^{2k}}{1-q^{2k}} - \sum_{k \text{ odd}} \frac{kq^{2k}}{1-q^{2k}} = \sum_{k \text{ even}} \frac{kq^{2k}}{1-q^{2k}} = \sum_{n \geq 1} (2n-1)u_{4n-2}
\]

as required.

The third identity can be shown by writing

\[ u_l u_{k-l} = \sum_{m \geq 1} \sum_{n \geq 1} q^{ml} q^{n(k-l)} \]

\[ = \sum_{n=m} q^{ml} q^{n(k-l)} + \sum_{m<n} q^{ml} q^{n(k-l)} + \sum_{m>n} q^{ml} q^{n(k-l)} \]
and noting that the three summands on the right hand side are $u_k, u_k u_l$ and $u_k u_{k+l}$. The last fact follows from the third by replacing $k$ by $k + l$. q.e.d.

We define the power series

$$\theta(q) = \sum_{n \in \mathbb{Z}} q^{n^2} = 1 + 2q + 2q^4 + 2q^9 + \ldots$$

This function is known as Jacobi’s theta function. Let $r_k(n)$ be the number of ways of writing $n$ as a sum of $k$ squares. By definition of the theta function we have $\theta(q) = \sum_{n=0}^{\infty} r_1(n) q^n$. Moreover, since $q^n q^m = q^{n+m}$ it is easy to verify that as power series we have

$$\theta^k(q) = \sum_{n=0}^{\infty} r_k(n) q^n$$

We would like to compute $r_4(n)$. We will always assume that $|q| < 1$. It can be easily verified that this condition will guarantee the absolute convergence of all the series that will interest us in the following arguments. Consequently we will obtain actual identities of values rather than of formal power series. Consider a series of the form

$$f(q) = \sum_{n=1}^{\infty} a_n u_n(q)$$

Formally we have

$$f(q) = \sum_{d=1}^{\infty} a_d u_d(q) = \sum_{d=1}^{\infty} a_d \left( \sum_{j=1}^{\infty} q^{jd} \right) = \sum_{n=1}^{\infty} b_n q^n$$

with

$$b_n = \sum_{d|n} a_d$$

We apply this formula with $a_d = \chi(d)$, where $\chi$ is the function introduced above, $\chi(d) = (-1)^{(d-1)/2}$ for $d$ odd and zero otherwise. Following theorem 43 we obtain

$$\theta^2(q) = \sum_{n=1}^{\infty} r_2(n) q^n = 4 \left( \sum_{m=1}^{\infty} \chi(m) u_m(q) \right) = 4(u_1 - u_3 + u_5 - u_7 + \ldots)$$

Recall that $\chi(d) = \sin\left(\frac{\pi d}{2}\right)$. We can similarly define the period 4 function $\psi(d) = \cos\left(\frac{\pi d}{2}\right)$. The following two identities relating $\chi$ and $\psi$ can be proved either as particular cases of basic trigonometric identities or simply by checking all 16 possible cases $0 \leq m, n \leq 3$,

$$\frac{1}{2} \chi(m) = \frac{1 + \psi(m)}{2} + \sum_{j=1}^{m-1} \psi(j)$$

and

$$2\chi(m)\chi(n) = \psi(m-n) - \psi(m+n)$$
Having gone through these preliminaries we come to the main computation. We calculate
\[
\left( \frac{1}{4} + \sum_{n \geq 1} u_n \chi(n) \right)^2
\]
which by (224) is simply \( \theta^4/16 \). Using (225) and (226) we get
\[
\left( \frac{1}{4} + \sum_{n \geq 1} u_n \chi(n) \right)^2 = \frac{1}{16} + \frac{1}{2} \sum_{n \geq 1} u_n \chi(n) + \sum_{m,n \geq 1} u_n u_m \chi(n) \chi(m)
\]
(227)
\[
= \frac{1}{16} + \frac{1}{2} \sum_{n \geq 1} u_n \left( \frac{1 + \psi(m)}{2} + \sum_{j=1}^{m-1} \psi(j) \right) + \frac{1}{2} \sum_{m,n \geq 1} u_n u_m (\psi(m-n) - \psi(m+n))
\]
(228)
We rearrange the summands according to the term \( \psi(k) \) which appears in them, grouping together summands with \( \psi(k) \) together with summands with \( \psi(-k) = \psi(k) \). Consequently, we can write
\[
\left( \frac{1}{4} + \sum_{n \geq 1} u_n \chi(n) \right)^2 = \frac{1}{16} + P_0 + \sum_{k \geq 1} P_k \psi(k)
\]
where the \( P_k \) are series.

The term \( P_0 \) consists of the series \( \frac{1}{2} \sum_{n \geq 1} u_n \) coming from the second summand of (228) and a series \( \frac{1}{2} \sum_{n \geq 1} u_n^2 \) coming from the third summand, when taking \( n = m \). Applying identity (1) from the lemma we see that
\[
P_0 = \frac{1}{2} \sum_{n \geq 1} n u_n
\]
For \( P_k, k > 0 \), we get a contribution of
\[
\frac{1}{2} u_k \sum_{n \geq k+1} u_n = \frac{1}{2} u_k + \sum_{l \geq 1} u_{k+l}
\]
from the second summand in (228) and a contribution of
\[
\frac{1}{2} \sum_{m=n-k} u_m u_n + \frac{1}{2} \sum_{n-m=k} u_m u_n - \frac{1}{2} \sum_{m+n} = ku_m u_n
\]
coming from the third summand of (228). Using identities (3) and (4) from the lemma we obtain
\[
P_k = u_k \left( \frac{1}{2} + \sum_{l \geq 1} (u_l - u_{k+l}) \right) - \frac{1}{2} \sum_{l=1}^{k-1} (1 + u_l + u_{k-l})
\]
\[
= u_k \left( \frac{1}{2} - \frac{1}{2} (k - 1) \right) + \sum_{l=1}^{k} u_l - \sum_{l=1}^{k-1} u_l = u_k (1 + u_k - \frac{1}{2} k)
\]
Putting it all together, plugging in the values of $\psi$ and applying identity (2) from the lemma we get

$$
\left( \frac{1}{4} + \sum_{n \geq 1} u_n \chi(n) \right)^2 = \frac{1}{16} + \frac{1}{2} \sum_{n \geq 1} nu_n + \sum_{k \geq 1} u_k (1 + u_k - \frac{1}{2} k) \psi(k)
$$

$$
= \frac{1}{16} + \sum_{k \geq 1} u_k (1 + u_k) \psi(k) + \frac{1}{2} \sum_{k \geq 1} ku_k (1 - \psi(k))
$$

$$
= \frac{1}{16} - \sum_{m \geq 1} (-1)^{m-1} u_{2m} (1 + u_{2m}) + \frac{1}{2} \sum_{m \geq 1} (2m-1) u_{2m-1} + 2 \sum_{m \geq 1} (2m-1) u_{4m-2}
$$

$$
= \frac{1}{16} - \sum_{m \geq 1} (2m - 1) u_{4m-2} + \frac{1}{2} \sum_{m \geq 1} (2m-1) u_{2m-1} + 2 \sum_{m \geq 1} (2m - 1) u_{4m-2}
$$

$$
= \frac{1}{16} + \frac{1}{2} \sum_{m'} m' u_{m'}
$$

where the last sum over $m'$ refers to all values of $m$ which are not divisible by 4. We conclude that

$$
\theta^4 = 1 + 8 \sum_{m'} m' u_{m'}
$$

and using the definition of $u_m$ we obtain

$$
\theta^4(q) = 1 + 8 \sum_{m'} m' \sum_{j \geq 1} q^{m'j} = 1 + 8 \sum_{n \geq 1} (\sum_{m'} m') q^n
$$

Thus, the number of representations of $n$ as a sum of 4 squares is 8 times, the sum of divisors which are not divisible by 4. In particular if $n$ is odd, then the number of representations is $8 \sum_{d|n} d$ as required. q.e.d.

While the proof above is "elementary", it is obvious that whoever devised it was a master manipulator of series identities. It is perhaps not surprising that the proof is due to Ramanujan, who indeed was a master manipulator of series. The proof here follows very closely the account in [89], theorem 385. More "conceptual" proofs can be found using the theory of modular forms which is discussed in the next chapter, see for example, [62].
Bibliography


[58] L. Euler, Variae observationes circa series infinitas (Various Observations about Infinite Series), published by St Petersburg Academy at 1737.


[77] L. Gerritzen and M. van der Put, Schottky groups and Mumford curves, Lecture Notes in Mathematics (LNM) 817, Speringer Verlag, 1980.
[114] V. Kolyvagin, Finiteness of E(Q) and X(E,Q) for a class of Weil curves, Math. USSR, Izv., 32, 523541, 1989.
[121] M. Lowe and F. Merkl, Moderate deviations for longest increasing subsequences: The upper tail, Communications in pure and applied mathematics, 54 (2001), 1488-1520.


Index

$q$-normal, 38
Action, 89
Adjacency matrix, 214
Airplane Boarding problem, 54
Airplane boarding problem, 55, 57
Back-to-front boarding, 58
Blocking, 56
Boarding partial order, 56
Boarding time, 59
Analytic function, 213
Anti-chain, 48
Arzela-Ascoli theorem, 201
Atlas, 68
Basis of a topology, 197
Beltrami equation, 91
Bias model, 29
$I$-model, 30
PQRS model, 29
Self similar, 29
Blocking potential, 87
Boarding metric, 88
Borel-Cantelli lemma, 202
Boundary, 197
Bounded Pareto distribution, 160
Branching process, 112
Cache algorithm
   Static, 34
Caching algorithm, 32
Causal partial order, 71
Causal relation, 71
Causal set, 74
Causally convex, 71
Causally convex (Lorentzian geometry), 64
Cayley graphs, 134, 215
Chain, 48
Chart, 68
Chebyshev’s inequality, 205
Closed set, 197
Closure of a set, 197
Conductor
   Of a character, 209
Configuration graph, 110
Congestion parameter, 88
Connected
   Component, 214
   Graph, 214
Conserved quantity, 91
Constant of motion, 91
Curve, 64
   Parametrized, 64
   Causal, 64, 71
   Order preserving, 64
Cut metric, 18
Cycle, 214
Cycle threshold, 111
Degree (of a graph), 214
Density
   $k$-finite, 151
Diameter, 73
Disk scheduling problem, 47, 57
   Service time, 49
   Service time random variable, 60
   The ABZ algorithm, 51
   Tour, 49
Distance (in a graph), 214
Double Pareto distribution, 161
Energy, 91
Entropy, 11
   Conditional, 33
   joint, 33
   Spatial, 34
   Temporal, 34
Equi-continuity, 200
Equi-distributed sequence, 38
Erdos-Renyi model, 111
Euclidean metric, 14
Euler-Lagrange equation, 90
Forest, 214
Fourier transform, 210
Functional, 89
Fundamental domain, 164
Future (poset), 64
Gaussian integers, 117
Generating function, 206
Geodesic, 104
Girth, 214
Graph, 214
Acyclic, 214
Edges, 214
regular, 214
Vertices, 214
Hamiltonian, 91
Hamiltonian integers, 117
Hamiltonian rational quaternions, 117
Hausdorff metric, 73
Hecke involution, 142
Hit ratio, 32
holomorphic function, 213
Homothety of lattices, 165
Hit ratio, 32
Horizontal partial order, 49
Hyper-metric, 19
i.i.d. sequence, 205
Independent Reference Model (IRM), 10
Interval
Open (Lorentzian geometry), 64
Interval (poset), 64
Job size, 138
Job-size distribution, 139
Kolmogorov-Sinai entropy, 44
Lattice, 165
Leaf, 214
Length, 72
Of a curve, 66
Of a walk, 214
Level
Of a density, 159
Of a Hecke involution, 142
LIFO, 154
Linearly coupled servers, 187
LJF, 154
Log-normal distribution, 161
Lorentz transformation, 62
Lorentzian geometry, 60
Lorentzian manifold, see also
Space-time manifold, 70
Lorentzian metric, 69
Manifold, 68
Marginal measure, 29
Markov chain, 207
Homogeneous, 207
Irreducible, 207
Stationary, 207
Markov's inequality, 205
Maximal curve, 73
Mellin transform
Incomplete, 144
Mellin transform, 144
Meromorphic function, 213
Metric, 13
$\ell_p$, 18
Metric space, 13
Mini-market problem, 138
Minkowski
Distance, 65
Minkowski space, 60, 61
Causal partial order, 62
Matrix, 61
metric, 61
Vector classification, 61
Model
Super-additive, 16
Moment of a random variable, 205
Mutual information, 34
Nice domain, 74, 106
Octonions, 134
Open set, 197
Orientation (Of a graph), 215
Orthogonal, 67
Pareto distributions, 160
Partially ordered set (poset), 47
Past (poset), 64
Patience sorting, 53
Poisson arrivals, 138
Poset
Linearly ordered, 48
Comparable elements, 48
Incomparable elements, 48
Minimal element, 48
Positive partial order, 53
Primitive (character), 209
Quadratic residue, 116
Quaternion
  Conjugate, 117
  Irreducible, 120
  Norm, 117
  Right divisor, 118
  Right g.c.d., 118
  Right sided relatively prime, 119
  Unit, 117
Queueing theory, 137

Random boarding policy, 58
Random variable, 205
Read hit, 32
Relative entropy, 11
Riemann zeta function, 143
Rooted tree, 214

Scheduling theory, 137
Seek time, 8
  function, 8
Self dual density, 159
Semi-metric, 13
Servers
  Coupled, 138
  Identical, 138
  Linearly coupled, 138
  Speed, 138
SITA policy, 153
SJF, 154
Space-time, 70
  Causal, 71
  Discrete, 74
  Strongly causal, 71
Space-time manifold, 68
Square Euclidean metric, 14
Squared coefficient of variation, 151
Standard deviation, 205
Standard Minkowski space, 61
State
  Ergodic, 207
  Null, 207
  Periodic, 207
  Persistent, 207
  Return time, 207
state
  Transient, 207
State (of a Markov chain), 207
Stationary distribution, 208
Support
  Of a set of files, 110
Time orientation, 70
  Of Minkowski space, 62
  Time parameter, 65
Topological space, 197
Topology, 197
  \textit{C}^0, 73
Trace, 9, 28
Tree, 214
Unicycle, 214
Uniform convergence, 200
Utilization, 139

Variance, 205
Vector
  Causal, 61, 70
  Light-like, 61, 70
  Space-like, 61, 70
  Tangent, 70
  Time-like, 61, 70
Vector field, 70
vertex transitive, 135
Vertical partial order, 50
Volume, 67, 73
Walk (in a graph), 214
Wave propagation process, 48
  Wave front, 48
Weight, 142