PRIME NUMBER PATTERNS

KRISTAL KOH MEI LING

PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF BACHELOR OF SCIENCE WITH HONOURS

MATHEMATICS WITH ECONOMICS PROGRAM
SCHOOL OF SCIENCE AND TECHNOLOGY
UNIVERSITI MALAYSIA SABAH

April 2008

UMS
UNIVERSITI MALAYSIA SABAH

PUMS99:1

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS@

JUDUL: PRIME NUMBER PATTERNS

IJAZAH: B.Sc (Hons)

SAYA KRISTAL KOH MEI LING SESI PENGAJIAN: 2007/2008
(HURUF BESAR)

mengaku membenarkan tesis (LPSM/Sarjana/Doktor Falsafah) ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis adalah hakmilik Universiti Malaysia Sabah.
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/)

[ ] SULIT (Mengandungi maklumat yang berdarjah keselamatan atau Kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

[ ] TERHAD (Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

[✓] TIDAK TERHAD

Disahkan Oleh

(TANDATANGAN PENULIS) (TANDATANGAN PUSTAKAWAN)

Alamat Tetap: 34, JALAN ANGSANA 9, BANDAR BOTANIK, 41200, KLANG, SELANGOR DARUL EHSAN.

Nama Penyelia: Tiong Kun ...

Tarikh: 24/4/2008 Tarikh: 22/04/2008

CATATAN:- *Potong yang tidak berkenaan.
**Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa /organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis ini perlu dikelaskan sebagai SULIT dan TERHAD.
@Tesis dimaksudkan sebagai tesis bagi Ijazah Doktor Falsafah dan Sarjana secara penyelidikan atau disertai bagi pengajian secara kerja kursus dan Laporan Projek Sarjana Muda (LPSM).

UMS UNIVERSITI MALAYSIA SABAH

DECLARATION

I hereby declare that this dissertation contains my original research work. Sources of findings reviewed herein have been appropriately acknowledged.

18 April 2008

KRISTAL KOH MEI LING

HS2005-2651

UMS
UNIVERSITI MALAYSIA SABAH

CERTIFIED BY

Signature

1. **SUPERVISOR**

   (Mr. Tiong Kung Ming)

2. **EXAMINER 1**

   (Dr. Aini Janteng)

3. **EXAMINER 2**

   (Cik Khadizah Ghazali)

4. **DEAN**

   (Supt./Ks. Prof. Madya Dr. Shariff AK Omang)

UMS
UNIVERSITI MALAYSIA SABAH

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor, Mr Victor Tiong Kung Ming, for his infinite guidance and motivation to the completion of this dissertation, for the countless hours and effort spent in providing invaluable counsel and insight.

I would also like to take this opportunity to thank my family for their continual support and prayers, being a source of encouragement all the time.

Above all, I want to thank God for His abundant grace and inspiration. Without Him, truly this dissertation would have never been possible.

UMS
UNIVERSITI MALAYSIA SABAH

ABSTRACT

This dissertation focuses on studying the patterns in the prime number distribution. It visualizes the distribution of prime numbers in a graphic manner through the usage of the Microsoft Excel software. This study is restricted to the first 20,000 primes. A list of the prime numbers is first obtained from Project Gutenberg online. By performing several methods of transformations, which are gaps between the first and second numbers, logarithm of the numbers, progressive ratio, progressive mean and progressive standard deviation, the transformed data are graphed in forms of scatter charts and radar charts. These charts are analyzed and compared with control data, which are number series that are increasing because prime numbers is an increasing number series. Prime numbers have been known to appear randomly, however, through the study of the graphs, it shows some regularity. This regularity is shown when a radar plot of primes, its progressive means and progressive standard deviation display a shell-like formation. All in all, although prime numbers seem to be scattered and occur in a somewhat random form, this dissertation shows that there are more underlying patterns with regularity that have not been totally discovered.

UMS
UNIVERSITI MALAYSIA SABAH

POLA NOMBOR PERDANA

ABSTRAK

Disertasi ini berpusatkan kajian mengenai pola nombor perdana. Ia menggambarkan pola nombor perdana dalam bentuk graf menggunakan Microsoft Excel. Analisis ini dihadkan kepada 20,000 nombor perdana yang pertama. Pada mulanya, senarai nombor ini diperoleh dari laman web Project Gutenberg. Dengan mengambil kaedah-kaedah tranformasi yang berbeza, iaitu jarak antara nombor pertama dan kedua, logaritma nombor, nisbah progresif, min progresif dan sisihan piawai progresif, data yang telah mengalami transformasi ini dilakarkan dalam bentuk graf 'scatter' dan graf radar. Graf-graf ini ditafsir dan dibandingkan dengan beberapa set kawalan. Set kawalan adalah siri nombor yang meningkat kerana nombor perdana merupakan siri nombor yang meningkat. Diketahui bahawa nombor perdana wujud secara rawak. Walaubagaimanapun, melalui kajian graf-graf, ia menunjukkan sesuatu pola. Daripada graf radar nombor perdana, min progresif serta sisihan piawai progresifnya, wujud suatu pola berbentuk kerang. Secara keseluruhan, sungguhpun susunan nombor perdana kelihatan rawak, namun disertasi ini menunjukkan terdapat pola memaparkan keseragaman yang belum ditemui lagi.

UMS
UNIVERSITI MALAYSIA SABAH

LIST OF CONTENTS

UMS
UNIVERSITI MALAYSIA SABAH

UMS
UNIVERSITI MALAYSIA SABAH

UMS
UNIVERSITI MALAYSIA SABAH

UMS
UNIVERSITI MALAYSIA SABAH

LIST OF FIGURES

UMS
UNIVERSITI MALAYSIA SABAH

UMS
UNIVERSITI MALAYSIA SABAH

CHAPTER 1

INTRODUCTION

1.1 Introduction

Prime numbers occupy a field in the mathematics that is relatively new, in a sense that not many discoveries have been made of recent. There is still much room for advancement that requires further study and research.

1.1.1 Definition

Prime numbers are said to be the building block of all numbers, or the basic "atoms" in integer systems (Jameson, 2003). Mollin (1998) calls prime numbers the building bricks of arithmetic. Prime numbers are also known as primes. Primes are defined when

$$p \in N, (p < 1)$$

given that p does not have any positive divisors, other than p and 1, then p is a prime. Any integer $p > 1$ is known as a prime if its only positive divisors are 1 and itself (Jones and Jones, 1998). Prime numbers are also numbers that do not resolve into smaller factors (Hardy and Cunningham, 1932). However, if $n \in N$, (n > 1) and n is

UMS
UNIVERSITI MALAYSIA SABAH

not a prime, then n is called a composite. Composites are also integers that are products of smaller integers (Robbins, 2006). Basically, any non-prime positive integer greater than 1 is a composite (Schumer, 1996). Definition 11 in Euclid's Elements in Book VII states that a prime number is that which is only measured by a unit. In other words, a prime number can only be divided by itself and 1, and leaves no remainder.

1.1.2 Distribution of Prime Numbers

At this moment in time, what is known is that, there are infinite amount of primes (Feferman, 1964). The first impression of a sequence of primes is one of utmost irregularity. Yet somehow the primes seem to wear off. Progressing down the set of natural numbers, the numbers of primes tend to be more sparse (Brown, 1978). This means that the primes decrease with a somewhat regularity. However, it does not mean that there will be finite amount of primes, since Euclid has proven its infinitude (Grosswald, 1984).

In detail, its distribution still seems to be utmost irregular. In an interval of many primes, there may appear long progressions of composite numbers consecutively (Nagell, 1981). One can only determine a prime number if it is not too large. Also, regrettably, there is still no easy formula to calculate or observe the location of the prime numbers (Ribenboim, 2000). The unpredictable gaps, stretches and openings between primes make finding a general formula that can encapsulate all prime numbers a daunting process.

UMS
UNIVERSITI MALAYSIA SABAH

1.1.3 Gap Between Prime Numbers

Some mathematicians have also studied prime distribution by observing the gap between the primes. A prime gap is the interval bordered by two consecutive primes (Nyman and Nicely, 2003). A new discovery by Szpiro (2003) reveals the possible presence of some structure when considering the gaps between the gaps and the higher-order counterparts. He made tests on frequency distributions of the gaps for different orders. First-order gaps were grouped together in multiples of six, and he supported his claim by comparing this series to a series of scrambled data that he simulated. However, this finding is recent and no extra studies have been undertaken to prove his case yet.

1.1.4 Prime Numbers Today

Mathematics is known as a research of patterns. Although mathematicians have been studying prime numbers for thousands of years, yet there are more open problems today than before (Conway and Guy, 1996). The vain search for a formula that yields all the prime numbers have perplexed mathematicians for centuries (Brown, 1978). When one is able to comprehend the pattern of the primes, one might be fully able to grasp the entire system of mathematics, since primes are the building blocks of all numbers (Sautoy, 2003).

Prime numbers also play a vital role in the fields of physics and biology. One new finding is that the quantum energy levels of a particle can be charted onto a sequence of prime numbers (Kumar, Ivanov and Stanley, 2007). Another was that

UMS
UNIVERSITI MALAYSIA SABAH

some congruence was found between the primes' distribution, elemental periodicity and the DNA code (Boeyens, 2003).

Today, prime numbers are used widely in modern scientific applications, especially in the field of cryptology (Maurer, 2001). Cryptography is the design and application of information protection techniques (Niederreiter, 2002). The field of cryptography includes methods needed to protect data (Trappe and Washington, 2006). The security of public-key systems is based on the difficulty of factorizing large prime numbers (Maurer, 1991). Encryption is the act of modifying files using a secret code so as to be unintelligible to unauthorized parties. One simple example is the vast use of bank cards, whereby the card is protected by a secret access code. Another is when making any transaction on the internet, say a credit card purchase, the safekeeping of that transaction is made certain by the use of cryptography. Some systems operate on the product of two primes to do the coding, but instead needs the prime itself to crack the code. It is tough to find the factors of the number, as they are usually about 100 digits long. Even by using the super computer, it will take ages to find the factors. Hence, since primes are not fully understood, security of the encryption is still not totally breached (Schumer, 1996). However, the full understanding of primes will prove fundamental to the advancement of the internet, technology, science and mathematics as a whole.

1.2 Objectives Of Research

The objectives of this research:

i. To study prime number distribution using experimental modeling to produce transformations of the prime numbers. By running considerable

UMS
UNIVERSITI MALAYSIA SABAH

number of tests, multiple possibilities are explored and analyzed in discovering new distributions of prime numbers.

ii. To search for and to justify significant patterns in the transformed sets of data. To support these justifications, the transformed sets of prime number data are compared with various other integer sequences that are used as control sets.

iii. To attempt to visualize prime number distribution by utilizing Microsoft Excel. This research hopes to shed light on the prime number distribution in a more graphic and comprehendible way, such that it would be simpler to envision and imagine the distributions.

iv. To come up with propositions and suggestions on the possible distribution of prime numbers.

1.3 Scope Of Research

To restrict the scope of this research, a few constraints will be included:

i. Although there has been many prime numbers that have been discovered, all experiments and testing in this research will be confined to a maximum of 20,000 primes only. The complete data of 20,000 primes is obtained from one of the sites of Project Gutenberg and cross-checked with other computer software.

UMS
UNIVERSITI MALAYSIA SABAH

ii. The experiments will run on sets of data of 500, 1,000, 5,000, and 10,000 unless otherwise stated.

iii. Visual representations will be mainly depicted through one software, that is Microsoft Excel 2003. The graphs chosen to represent results of this research are the scatter chart and radar chart.

1.4 Significance of Research

No other mathematical unit has so captured minds more than prime numbers. Instead of merely toying around and staring at formulas, example Euler's formula $n^2 + n + 41$, this study aims to bring the mind to visualize the distribution of prime numbers. By looking at visual representations, one can picture more vividly how the primes are dispersed, for example the pioneering work of Stanislaw Ulam (Brown, 1978). Since the study of prime numbers has streamlined far and wide into different majors previously discussed like physics, biology and cryptology, this research aspires to renew some ideas about the distribution of primes that are significant and worthy of further exploration.

UMS
UNIVERSITI MALAYSIA SABAH

CHAPTER 2

LITERATURE REVIEW

2.1 History

The first records of primes came from ancient Greeks (Grosswald, 1984), though it is believed that ancient Egyptians learnt of primes first. The Greek mathematicians at Pythagoras' school studied about perfect and amicable numbers between 500B.C. to 300B.C. The book by Euclid the Greek in 300BC, Euclid's Elements, is well known for providing the oldest known proof about the infinitude of primes. Euclid's Elements also contains some vital theorems about primes (Jones and Jones, 1998).

2.1.1 Mathematicians

This segment focuses on the mathematicians and people who made significant contributions to the discovery of prime numbers and its distributions.

UMS
UNIVERSITI MALAYSIA SABAH

a. Pierre de Fermat

French Pierre de Fermat, a judge, appeared around year 1937. He first proved prime numbers in the form

$$4n+1$$

can be written as the total of two squares and was also able to show how any number could be written as a total of four squares. This he proved by using the method of infinite descent (Feferman, 1964). Then Fermat stated a theorem which he wrote in his Bachet's Latin translation of the Arithmetica of Diophantos (Jones and Jones, 1998). If p is prime then

$$a^p = a \bmod p$$

where a is any integer. This theorem is later known as Fermat's Little Theorem.

b. Marin Mersenne

Marin Mersenne came up with the formula

$$M_n = 2^n - 1$$

known as the Mersenne prime, (Schumer, 1996). Most numbers of this form with n prime are prime, but not all. He claimed that $2^n - 1$ was prime for many numbers, but he gave no evidence to support his claim. In year 1588, Cataldi proved M_{19} to be the largest prime. This record was held for about 200 years until Leonard Euler from Sweden came about. Euler was able to prove in 1772 that M_{31} is a prime, and he held this record for another century before other people uncovered more primes.

UMS
UNIVERSITI MALAYSIA SABAH

c. Leonard Euler

Leonard Euler from Sweden devised an important function known as the zeta function

$$\varsigma(s)=\sum_{n=1}^{\infty}\frac{1}{n^{s}}=\prod_{p}^{\infty}\frac{1}{1-p^{-s}}$$

as an infinite product concerning the prime numbers. This formula expresses the sole factorization of integers as product of the prime numbers (Ribenboim, 1996). Euler proved that the sum of prime numbers' inverses is divergent. This formula will be expanded on in Section 2.1.1e. In attempting to find a formula that best describes the prime number distribution, Euler then came up with the formula

$$n^2+n+41$$

This formula is unique in a sense that, it only works from n = 0 till 39. At n = 40 onwards, this formula will fail to generate prime numbers. However, an amazing discovery is made when numbers from 41 to 440 are placed in a square spiral, starting with number 41 in the middle and turning anti-clockwise. The numbers on the diagonal shown in Figure 2.1 are known as Euler's formula primes. From further examination, many other prime numbers that are not found by Euler's formula also tend up on diagonals (Corliss, 1985). This will be further expounded on in Section 2.2.4.

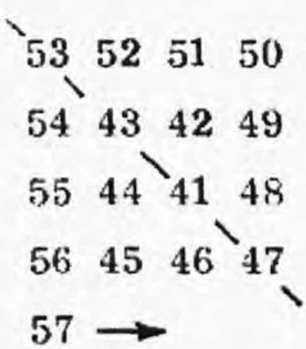

Figure 2.1 Euler's formula primes (Corliss, 1985).

UMS
UNIVERSITI MALAYSIA SABAH

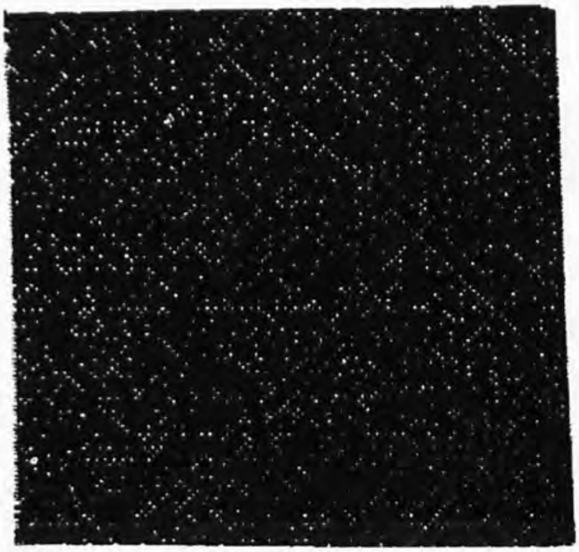

Figure 2.2 Euler's formula primes in a 20 x 20 grid (Corliss, 1985).

d. Carl Friedrich Gauss

Carl Friedrich Gauss of Germany emerged in the year 1777 to 1855. At age fifteen, he came up with his own mathematical masterpiece, Disquitiones Arithmeticae, which was published in year 1801 (Mollin, 1998). The launch of Gauss into the prime number research began after he was given a table of prime numbers. At his point of entry, the study on prime number distribution has already stagnant for a long period. Instead of asking the norm question of which numbers are prime, he asked instead, how many primes should there be in a given range of numbers.

Gauss made a guess that primes was chosen with a dice. He came up with a table with n in one column, and $\pi(n)$ in another. $\pi(n)$ represents the number of primes up to n. In the third column, he calculated the gap between the primes. Gauss noticed a pattern. Each time n is multiplied with 10, the numbers in the third column add on by about 2.3. He then converted this pattern into logarithmic form which is

PERPUSTAKAAN UNIVERSITI MALAYSIA SABAH

UMS UNIVERSITI MALAYSIA SABAH

$$\log_e 10 = 2.3026$$

By doing so, he predicts the number of primes less than *n*. As *n* gets bigger, log *n* increases as well. Hence the chance of getting a prime number reduces with increasing *n*. This implies that the primes appear lesser and lesser progressing along the natural number system. Then using the model of a prime number dice, Gauss projected the number of primes by

$$\frac{1}{\log 2} + \frac{1}{\log 3} + ... + \frac{1}{\log n}$$

This projection is known as Gauss' logarithmic integral, Li(*n*), which stands for the average distance between consecutive prime numbers (Szpiro, 2003). To test the accuracy of Gauss' projection, a graph of Li(*n*) is plot against the real number of primes.

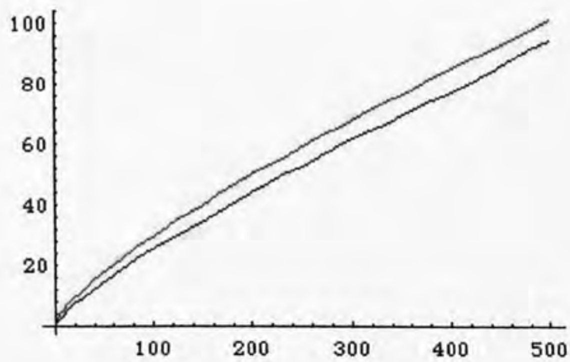

Figure 2.3 Li(*n*) against primes.

The line above shows Gauss' guess whereas the line below is the actual plot of prime numbers. His projection is not exactly accurate, yet as n gets bigger, the percentage error gets smaller. Gauss has revealed the "Prime Number Dice" that nature used for selection of primes, where the number of sides grew like the logarithmic function. The trouble Gauss had was finding out how this dice landed. Also, he came up with Gauss'

UMS
UNIVERSITI MALAYSIA SABAH

REFERENCES

Boeyens, J. C. A. 2003. Number Patterns in Nature. *Crystal Engineering* **6**, 167-185.

Brown, S. I. 1978. *Some Prime Comparisons.* The National Council of Teachers of Mathematics, United States of America.

Conway, J. H. and Guy, K. G. 1996. *The Book of Numbers.* Copernicus Springer-Verlag, New York.

Corliss, W. R. 1985. The Fabric of Prime Number Distribution. *Science Frontiers* **42,** 25.

Feferman, S. 1964. *The Number Systems: Foundations of Algebra and Analysis.* Addison-Wesley Publishing Company *Inc., London.*

Grosswald, E. 1984. *Topics from the Theory of Numbers.* 2nd edition. Birkhauser, Boston.

Hardy, G. H. and Cunningham, E. 1932. *No. 30: The Distribution of Prime Numbers.* Cambridge University Press, New York.

Jameson, G. J. O. 2003. *The Prime Number Theorem.* Cambridge University Press, United Kingdom.

UMS
UNIVERSITI MALAYSIA SABAH

Jones, G. A. and Jones, J. M. 1998. *Elementary Number Theory*. Springer-Verlag London Limited, Great Britain.

Kumar, P., Ivanov P. C. and Stanley, H. E. 2007. *Information Entropy and Correlations in Prime Numbers*. Boston University, Boston.

Lerner, K. L. 2007. *Cryptology and Number Theory*. Retrieved 17 September 2007 from http://www.espionageinfo.com/Cou-De/Cryptology-and-Number-Theory.html.

Maurer, U. M. 2001. *Generation of Prime Numbers*. Retrieved 17 September 2007 from http://www.crypto.ethz.ch/research/ntc/gpn/.

Maurer, U. M. 1991. Fast Generation of Prime Numbers and Secure Public-Key Cryptographic Parameters. *Journal of Cryptography* **8**, 123-155.

Mollin, R. A. 1998. *Fundamental Number Theory With Applications*. CRC Press LLC, United States of America.

Nagell, T. 1981. *Introduction to Number Theory,* 2nd edition. Chelsea Publishing Company, New York.

Niederreiter, H. 2002. *Coding Theory and Cryptology.* Singapore University Press and Scientific Publishing Co. Pte. Ltd, Singapore.

UMS
UNIVERSITI MALAYSIA SABAH

Nyman, B. and Nicely, T. R. 2003. New Prime Gaps Between 10^{15} and 5×10^{16}. *Journal of Integer Sequences* **6**.

Pomerance, C. 1985. The Quadratic Sieve Factoring Algorithm. *Lecture Notes in Computer Science* **209**, 169.

Project Gutenberg. 2001. *First 100,000 Prime Numbers*. Retrieved 18 July 2007 from http://www.gutenberg.org/dirs/etext93/prime12.txt.

Ribenboim, P. 1996. *The New Book of Prime Number Records.* Springer-Verlag London Limited, London.

Ribenboim, P. 2000. *My Numbers, My Friends: Popular Lectures on Number Theory*. Springer-Verlag London Limited, London.

Robbins, N. 2006. *Beginning Number Theory*, 2nd edition. Jones and Bartlett Publishers, United States of America.

Sautoy, M. 2003. *Why Are Primes So Important?* Retrieved 17 September 2007 from http://www.musicoftheprimes.com/why_are_important.htm.

Schumer, P. D. 1996. *Introduction to Number Theory*. PWS Publishing Company, United States of America.

UMS
UNIVERSITI MALAYSIA SABAH

Szpiro, G. G. 2003. The Gaps Between the Gaps: Some Patterns in the Prime Number Sequence. *Physica A* **341**, 607-617.

Trappe W. and Washington I. C. 2006. *Introduction to Cryptology with Coding Theory,* 2nd edition. Pearson Prentice Hall, New Jersey.

Weisstein, E. W. 2005. *Prime Spiral*. Retrieved 25 July 2007 from http://mathworld.wolfram.com/PrimeSpiral.html.

Weisstein, E. W. 2000. *Quadratic Sieve.* Retrieved 17 September 2007 from http://mathworld.wolfram.com/QuadraticSieve.html.

Wikipedia. 2007. *Great Internet Mersenne Prime Search.* Retrieved 18 July 2007 from http://en.wikipedia.org/wiki/GIMPS.

UMS
UNIVERSITI MALAYSIA SABAH