1

Introduction

1.1 Why This Study?

This book provides an account of the remarkable progress which has been made during the first two decades of the twenty-first century, reducing the size of the known gap length between infinite numbers of consecutive primes. Parts of the story, and the mathematicians who have been responsible for this progress, are well known. Here, in one book, are the mathematical details needed to follow the developments and, it is hoped, extend them. In addition, since computation is an important component of some of the methods which are used, a suite of Mathematica programs is provided for checking results and for further experimentation.

Most mathematicians believe there are an infinite number of twin primes. However, by 2013 it was not known for sure whether or not the minimum separation for pairs of consecutive primes tended to infinity. That is to say, if p_n is the nth prime with $p_1 = 2$, then

$$\liminf_{n\to\infty} p_{n+1} - p_n = \infty \ ?$$

In 2014, all this changed. Two groups and three individuals, Yitang Zhang, Polymath8a, James Maynard, Terence Tao and Polymath8b, both separately, together and progressively, first showed that $p_{n+1} - p_n \le 7 \times 10^7$ for an infinite number of p_n, and then, in many, many steps, lowered the upper bound to 246. This incredible progress on the question of prime gaps then ceased. As far as prime gaps is concerned, at the time of writing the last several years have not witnessed any further reductions in the proven minimal gap width.

It is the purpose of this book to explain how this progress came about. In particular, the book describes how it is rooted in the conjecture of Dickson, Hardy and Littlewood on finite patterns of primes, uses variations on the Selberg sieve and would not have happened at all were it not for the earlier breakthrough of Goldston, Motohashi, Pintz and Yildirim (known collectively as GMPY with the subset GPY also appearing quite often).

There is considerable technical detail and much computational support needed to derive the best results. Just because an idea has been used to derive a prime gap which is bigger than the current best, its deeper understanding and combination with other ideas and methods could bring about further improvements. In any case, the early ideas are an essential part of this story.

1.2 Summary of This Chapter

Section 1.3 gives an overview of the contents of the book. Section 1.4 describes Timothy Gowers' idea of a polymath project, and lists named contributors to Polymath8. Section 1.5 gives a timeline of the developments which are covered, in whole or in part, or which were a prelude to the breakthroughs. Section 1.6 discusses the twin primes constant and the Dickson–Hardy–Littlewood conjecture, which relates to the essential underlying concept of "admissible tuples" of integers. Section 1.7 delves into the nature of the prime gap distribution by discussing the issue of which prime gap up to increasing positive x is most common, and reports on the recent work on "jumping champions". Section 1.8 gives the derivation of some useful properties of the von Mangoldt function which are needed in the following sections and chapters. Section 1.9 is devoted to a discussion of the Bombieri–Vinogradov theorem, its statement, history and references to proofs. Section 1.10 introduces admissible tuples, which describe patterns of primes which are expected to repeat infinitely often, and the intriguing relationship between the Dickson–Hardy–Littlewood conjecture and the so-called second Hardy–Littlewood conjecture. Section 1.11 is a brief guide to the literature, including a film, some introductory secondary sources and the primary sources, with a reader's guide. This latter includes a note on which parts of the book could be skipped to get to the best result in minimum time.

In an end note, Section 1.12, in contrast to the small gaps between primes that are the principal focus of the book, successive results on large gaps between consecutive primes are summarized in a table with references.

1.3 History and Overview of These Developments

Searching for small gaps between consecutive primes is of course one way to approach the twin primes conjecture, which is one of the most celebrated unsolved problems in number theory. Some believe the conjecture was known to Euclid, but there is no evidence to support this. The first known mention is in Polignac in 1849 [170], where it is included in his more general conjecture that for each even integer $2k$ there is an infinite number of prime pairs (p,q), which do not need to be consecutive, such that $q = p + 2k$. A comment attributed to Dan Goldston from PBS's *Nova*:

No one really knows if Euclid made the twin primes conjecture. He does have a proof that there are infinitely many primes, and he or other Greeks could easily have thought of this problem, but the first published statement seems to be due to de Polignac in 1849 [170]. Strangely enough, the Goldbach conjecture, that every even number is a sum of two primes, seems less natural but was conjectured about 100 years before this.

Part of the context for this work was already established by Dickson in 1904 in [39]. He conjectures that infinite sets of primes occur in patterns, unless there is a modular condition outlawing a pattern. In detail, given a finite set of distinct integers $\mathscr{H} = \{h_1, h_2, \ldots, h_k\}$, then there exists an infinite number of translates $n + \mathscr{H}$ where all of the $n + h_j$ are prime, unless for some fixed prime p the size of $\mathscr{H}$ modulo p is p – in that case, for each n there is a j_n such that $p \mid n + h_{j_n}$ so for all n sufficiently large an element of the translated tuple cannot be prime. See Section 1.10 in this chapter. Twin primes have the pattern $\{0, 2\}$, and de Polignac's conjecture is the pattern $\{0, 2k\}$ for all $k \in \mathbb{N}$.

Hardy and Littlewood made a precise asymptotic form of Dickson's conjecture, which form is now called the Dickson–Hardy–Littwood conjecture, or DHL, and is described in Section 1.6. Work on this problem and on another conjecture Hardy and Littlewood made in the same paper [89], namely that for all $x, y \geq 2$ we have $\pi(x + y) \leq \pi(x) + \pi(y)$, added significantly to our knowledge of admissible k-tuples, i.e., those for which the number of distinct congruence classes modulo p is less than p for all primes p. See Section 1.10.3 in this chapter. It is an interesting fact, demonstrated in that section (but not used subsequently), that DHL and this second Hardy–Littlewood conjecture, cannot both be true at the same time.

The early work on gaps between primes focused on showing that there were gaps significantly smaller than the average. Up to real x, by the prime number theorem, this average is asymptotically $\log x$ of course. The first unconditional result is due to Erdős and was published in 1940. He proved that there were an infinite number of prime pairs strictly closer than this average by a small but unspecified amount, which was expected. His proof is given in Chapter 3, which includes also the reduction in gap size demonstrated by Bombieri and Davenport in 1966. Table 1.1 sets out most of the main other published work along these lines until 1988.

The breakthrough, signifying the start of the exciting modern era for work on the prime gap problem, was by Dan Goldston, Yoichi Motohashi, Janos Pintz and Cem Yildirim (known colloquially as GMPY), essentially done by 2006 but published in 2009. They showed that for any given $\epsilon > 0$ there are an infinite number of prime pairs $p < q \leq x$ with $(q - p) < \epsilon \log x$. See the introductory section of Chapter 4 for an overview of their work, or read Kannan Soundararajan's excellent paper [198]. Their principal tool was the sieve developed by Atle Selberg. Used for a wide variety of number theory problems, this tool was not only used in some of the early work (see for example Chapter 3), but also sufficiently flexible to be the main underpinning device for each of the later breakthroughs. See Chapter 2 for the elements of Selberg's method and an application which is used subsequently. There is also a brief overview of other commonly used sieves.

Then Yitang Zhang read the works of GPY, carefully and in detail, and thought long and hard about how to improve the methods. In particular, he found a way of improving the unconditional Bombieri–Vinogradov theorem, which is roughly speaking the generalized Riemann hypothesis for primes in arithmetic progressions on average, and used this improvement, and other very advanced techniques, to show that there were an infinite number of prime pairs $p < q$ such that $q - p \leq 7 \times 10^7$. Written in 2013, this work was published in 2014. His work was unusual, in that the results were based on explicit constants, but he made no claim that his upper bound was optimal. Indeed, in [215] we find a remark at the end of his section 1, which we paraphrase:

> This result is of course not optimal. To replace the constant 7×10^7 by a value as small as possible is an open problem that will not be discussed in this paper.

For some biographical background on Zhang, and an overview of his approach, see the introduction to Chapter 5. The reader could with benefit consult the wide-ranging survey of Andrew Granville [82].

It is important to stress that Zhang's work used multivariable exponential sum estimates based on the profound work of Deligne, solving the Riemann hypothesis for varieties over finite fields, and other techniques to derive his variation of Bombieri–Vinogradov's estimate. We don't cover most of these topics, but they should be treated fully elsewhere.

Under the leadership of Terence Tao, a Polymath question was formulated and group formed in 2013 (see some details in Section 1.4). Communicating over the Internet, this group and others began immediately to reduce the bound. A new concept, **densely divisible numbers** which generalize smooth integers, was formulated and extensive computations performed to optimize parameters, while remaining faithful to Zhang's approach. By the end of 2013, Polymath had reduced the bound to 4,680 (using a tuple of size 632), publishing this work in 2014 [173], with a more extensive write up on the archive [174], including extensive details of underlying computational methods. (There may be some doubt and confusion about the precise values of these bounds – at one stage, they were changing, both up and down, almost on a daily basis. For example, the author was able to reach a tuple size of 630 rather than 632 using Polymath's methods.)

Along the way, Polymath8a was able to reduce Zhang's bound significantly to 14,950 (using a tuple of size 1,783) without using Deligne, but with an estimate for exponential sums based on polynomials in a single variable, a so-called "Weil bound", which is much easier to derive and less reliant on very high-powered mathematical tools than Deligne's applications. This approach is included here in Chapter 8.

It is no longer news that this work was quickly superseded by James Maynard, who used a "mutivariable", or better to say a multidivisor, approach to Selberg's sieve, and some standard optimization over parameterized families of symmetric multinomials to reduce the bound to 600. Not only that, he completely sidestepped any use of an extended Bombieri–Vinogradov estimate or Deligne–Weil exponential sums. His work is the subject of Chapter 6 and was published in [138]. Of course, this discovery inspired another Polymath project under Terry Tao's leadership, and the bound was reduced to 246 with a tuple of size 50 by the end of 2014, with publication occurring [175] that year. The Polymath refinements of Maynard are described in Chapter 7.

The main part of the book concludes with Chapter 9, which gives a summary of some relatively recent (as of 2017) additional results proved using the methods of the principal authors using the Elliott–Halberstam conjecture, so these are conditional, or, for example, the best unconditional results on prime gaps, such as those of Maynard/Polymath. The appendices cover some of the more standard techniques which are employed by the authors, such as compact operators, Bessel functions, Stepanov's approach to the Weil bound for curves, some complex analysis needed for detailed estimates of the Riemann zeta function and an introduction to the dispersion method of Linnik. In addition, there is a minimanual for the suite of Mathematica programs, **PGpack**, which includes standard functions like Densely-DivisibleQ and VonMangoldt, but also Maynard and Polymath's algorithms, used in deriving their best results. There is also considerable support for Ignace Bogaert's Krylov method algorithm.

Finally, in this brief overview, it is good to keep in mind the dependencies of the work of the main authors on the work of many others. We represent some of these dependencies in the flow graph in Figure 1.1, adapted from that which was included in [150].

1.4 Polymath Projects and Members of Polymath8

A Polymath project played a significant role in lowering the proved gap between an infinite number of prime pairs. But what is a Polymath project and who were the named participants? In this section, an attempt will be made to answer these questions.

Polymath projects were initiated by Timothy Gowers in 2009. In the following I will summarize part of his thinking, which is extensively set out on Gowers' weblog "Is massively collaborative mathematics possible?", which enjoyed a very large amount of feedback. Within a few days, he had already formulated and

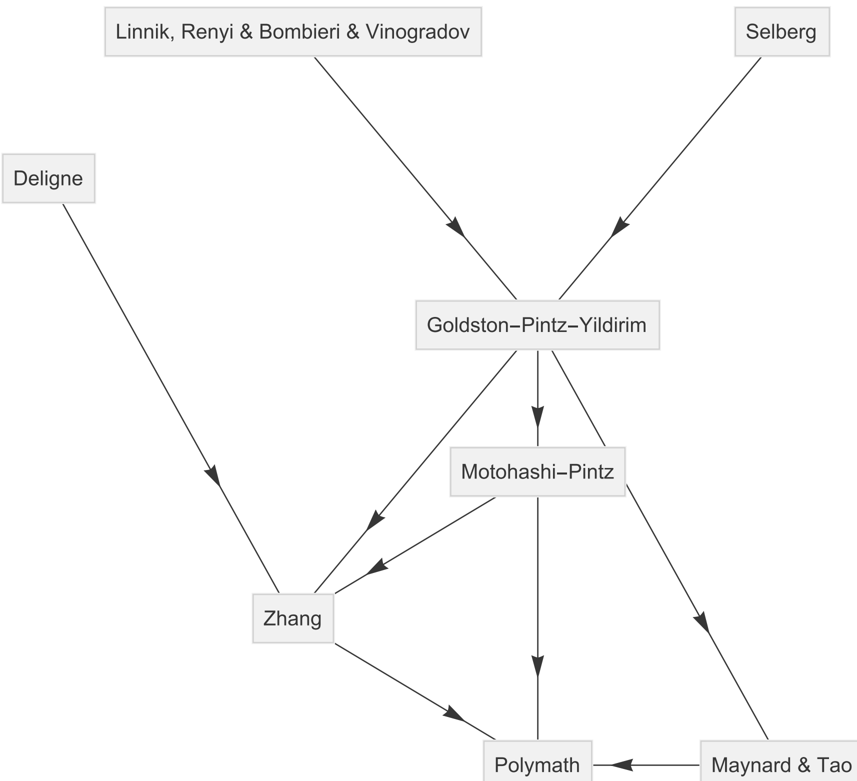

Figure 1.1 Some prime gaps developments dependencies.

revised a set of rules in "Questions of procedure", which described the process which was envisaged for Polymath projects. He had already listed several suitable topics and given some initial ideas, in "Background to a Polymath project", to start the process: the Hales–Jowett theorem, the Furstenberg–Katznelson theorem, Szemerédi's regularity lemma, the triangle removal lemma and so-called sparse regularity lemmas. By the time Tao had proposed Polymath8, seven other projects had been initiated.

Now Gowers, also known as Sir William Timothy Gowers, is a leading mathematician, having made many fundamental contributions to functional analysis, arithmetic combinatorics and graph theory. He was awarded the Fields Medal in 1998 for his research. Sitting on my bookshelf in pride of place is the *Princeton Companion to Mathematics* [80], a magnificent compendium for which he was the editor in chief.

Gowers' Concept of a Polymath Project

The idea of a Polymath project is anybody who had anything whatsoever to say about the project problem could contribute brief ideas, even if they were undeveloped and/or likely to be wrong. Some advantages:

• Sometimes luck is needed to have the idea that solves a problem. If lots of people think about a problem, then on probabilistic grounds there is more chance that one of them will have that bit of luck.

• Different people know different things, so the knowledge that a large group can bring to bear on a problem is significantly greater than the knowledge that one or two individuals will have. For example, consider a response, "The lemma you suggested trying to prove is known to be false." One can take weeks to discover this situation if one is on one's own.

• Different people have different characteristics when it comes to research. Some like to throw out ideas, others to criticize them, others to work out details, others to reexplain ideas in a different language, others to formulate different but related problems, others to step back from a big muddle of ideas and fashion some more coherent picture out of them and so on. Group members can specialize. In short, if a large group of mathematicians could connect their brains efficiently, they could perhaps solve suitable problems very efficiently.

• Why would anyone agree to share their ideas? Many work on problems in order to be able to publish solutions and get credit for them. And what if the Polymath collaboration resulted in a very good idea? Isn't there a danger that somebody would steal it from the Polymath to get the credit? Gowers rebutted these objections, but they are valid.

• Gowers doesn't believe that this approach is likely to be good for everything. Examples include the Riemann hypothesis in our current state of knowledge, or the solution of a very minor and specialized problem. He claimed there is a middle ground and gave examples. The first Polymath project, conceptualized and led by Gowers, was to find a new proof of the density Hales–Jewett theorem. (Hence the name D. H. J. Polymath as the author of some articles.) The statement is that for every $k > 0$ and $\epsilon > 0$, there is an $n \in \mathbb{N}$ such that for any subset A of words on $\{\{1,\ldots,k\} \cup \{x\}\}^n$ with $\#A > \epsilon k^n$, if $w(i)$ is the word obtained by replacing every occurrence of x by i, then A contains a set of the form, $\{w(i) : 1 \le i \le k\}$, a so-called "combinatorial line". The new proof came quickly [171] and showed that Gowers' concept was a positive contribution to mathematics research practice.

Polymath8 enjoyed a high level of participation by many mathematicians, some at the heart of the project and others more peripheral to the main developments, and they are not so easy to acknowledge. The following lists have been taken from the "Polymath8 grant acknowledgements" web page, accessible at the time of writing from `http://michaelneilsen.org/polymath1/index.php?title =Polymath8_grant_acknowledgements` (which should be entered into a search box as one expression with no spaces, returns or line-feeds):

Polymath8a authors or primary participants were Wouter Castryck, Etienne Fouvry, Gergely Harcos, Emmanuel Kowalski, Philippe Michel, Paul Nelson, Eytan Paldi, Janos Pintz, Andrew V. Sutherland, Terence Tao and Xiao-Feng Xie.

Polymath8b authors or primary participants were Ignace Bogaert, Aubrey de Grey, Gergely Harcos, Emmanuel Kowalski, Philippe Michel, James Maynard, Paul Nelson, Pace Nielsen, Eytan Paldi, Andrew V. Sutherland, Terence Tao and Xiao-Feng Xie.

Overviews of the work of Polymath8 have appeared in the *EMS Newsletter* in "The 'Bounded Gaps Between Primes' Polymath Project", by D. H. J. Polymath [172] and the *Notices of the AMS* in "Prime Numbers: A Much Needed Gap Is Finally Found", by J. Friedlander [59], and the cover description [25]. In addition, readers are strongly encouraged to first scan, and then consult from time to time as might be needed, the Polymath8 Home Page. This contains an impressive and expansive family of links and threads to the work of Polymath8 and others relating to prime gaps, including the latest results, write-ups, a timeline of bounds, code and data, errata, recent papers and notes, media reports and a bibliography. The home page is maintained by Michael Nielsen under the heading "Polymath1Wiki" with the title "Bounded Gaps Between Primes" and accessible at the time of writing from `http://michaelneilsen.org/polymath1/index.php?title=Main_Page`.

1.5 Timeline of Developments

This section is a brief but crucial element of this introductory chapter. The timeline of Table 1.1 shows how our knowledge of gaps between prime pairs has improved over almost 100 years, with improvements made by many leading mathematicians. It also shows how the work as published in regular journals splits into the period prior to 1988 and that between 2006 and 2015. This latter time range corresponds with the main contents of this book, but in Chapter 3 we do give some examples from the earlier period. The reader may wish to sometimes refer back to the table.

For all x sufficiently large, one can find primes in $[x, 2x]$ with $p_{n+1} - p_n$ less than the bound given in Table 1.1, and thus an infinite number of consecutive primes distant apart less than the given bound.

The dates attached to the bounds in Table 1.1 are not necessarily those of the first published form, which is often significantly earlier and often on the ArXiv preprint server, especially the dates for the work published during 2013 to 2015. The table is based on information in the online slides of Terence Tao, prepared for his lecture presented at the Latinos in the Mathematical Sciences Conference, UCLA, 2015.

Table 1.1 *Timeline of proved decreasing prime gaps.*

Year	Bound	People	Reference
1923	$(\frac{2}{3} + o(1)) \log x$	Hardy/Littlewood/GRH	[89]
1940	$(\frac{3}{5} + o(1)) \log x$	Rankin/GRH	[178, 179]
1940	$(1 - c + o(1)) \log x$	Erdős	[45]
1954	$(\frac{15}{16} + o(1)) \log x$	Ricci	[182]
1966	$(0.4665 + o(1)) \log x$	Bombieri/Davenport	[10]
1972	$(0.4571 + o(1)) \log x$	Pil'tjai	[164]
1973	$(0.4463 + o(1)) \log x$	Huxley	[109]
1975	$(0.4542 + o(1)) \log x$	Uchiyama	[209]
1977	$(0.4425 + o(1)) \log x$	Huxley	[110]
1984	$(0.4394 + o(1)) \log x$	Huxley	[111]
1988	$(0.2484 + o(1)) \log x$	Maier	[132]
2006	$o(\log x)$	Goldston/Motohashi/Pintz/Yildirim	[70]
2009	$C(\log x)^{\frac{1}{2}}(\log\log x)^2$	Goldston/Pintz/Yildirim	[72]
2013	7.0×10^7	Zhang	[215]
2013	$4,680$	Polymath8a	[26, 174]
2015	600	Maynard	[138]
2014	246	Polymath8b	[175]

1.6 Prime Patterns and the Hardy–Littlewood Conjecture

A study of the gaps between prime pairs is part of the more general search for patterns exhibited by finite sets of primes infinitely often. The ultimate strongly believed gap size is of course 2, and we begin this section with early approaches to the twin primes conjecture in a quantitative form. This is followed by the more general, and entirely relevant for the work of all contributors to the discoveries of this report, the so-called admissible k-tuples. Roughly speaking, these are sets of k distinct integers such that there are no modularity constraints which would prevent an infinite number of translates consisting of all primes. In a quantitative form, this gives rise to the Dickson–Hardy–Littlewood conjecture.

The **Hardy–Littlewood conjecture** of 1923 in particular gives the asymptotic relation that as $x \to \infty$

$$\pi_2(x) := \#\{p \le x : p + 2 \in \mathbb{P}\} \sim \eta C_2 \frac{x}{(\log x)^2} \text{ with } \eta = 2, \tag{1.1}$$

where the **twin primes constant** C_2 is the number

$$C_2 := \prod_{p \ge 3} \left(1 - \frac{1}{(p-1)^2}\right) = 0.6601618158468957\ldots.$$

Hardy and Littlewood derived their conjecture using Cramer's probability model, which is based on the assignment of a probability of $1/\log x$ to numbers in the range $[1, x]$ for x large, that they should be prime. If p is prime, the Cramer model gives a nonzero probability that $p+1$ should also be prime, but this is impossible for odd p. If the model held, then the probability that p and $p+2$ were prime would be $1/(\log x)^2$ leading to $x/(\log x)^2$ twin primes up to x. Hardy and Littlewood argued that the model could be used provided that it was corrected by a factor which is twice the twin primes constant.

Their argument was as follows: given a random integer n, then for both n and $n+2$ to be prime both must be odd, which occurs with probability $1/2$ rather than $1/4$ for a pair of random integers. Neither must be divisible by 3, which means we must have $n \equiv 1 \bmod 3$, which has probability $1/3$ rather than $4/9$ if they were independent and obeyed that constraint. If the prime is 5, then we must have $n \not\equiv 0, 3 \bmod 5$, giving a probability of $1-2/5$ rather than $(1-1/5)^2$. In general, for a prime $p \geq 3$ the probability neither n or $n+2$ is divisible by p would be $1-2/p$ rather than $(1-1/p)^2$ if they were independent. The factor 2 is a little different since both n and $n+2$ are odd with probability $1-1/2$ rather than $(1-1/2)^2$. Thus, to get the adjusted asymptotic count, we must multiply $x/\log^2 x$ by

$$\left(1-\frac{1}{2}\right)\left(1-\frac{1}{2}\right)^{-2}\prod_{p\geq 3}\left(1-\frac{2}{p}\right)\left(1-\frac{1}{p}\right)^{-2} = 2\prod_{p\geq 3}\left(1-\frac{1}{(p-1)^2}\right) = 2C_2.$$

Figure 1.2 is a plot of an exact count of the twin primes up to $n \in \mathbb{N}$, $T(n)$ times $(\log n)^2/n$, for $3 \leq n \leq 10^5$. Table 1.2 gives details of proved upper bounds for the leading coefficient η in (1.1), all for x sufficiently large.

Table 1.2 *Twin primes: upper bounds for the leading coefficient η.*

Year	Bound for η	People	Reference
1919	100	Brun	[24]
1966	$8+\epsilon$	Bombieri and Davenport	[10]
1983	$\frac{68}{9}+\epsilon$	Fouvry and Iwaniec	[55]
1984	$\frac{128}{17}$	Fouvry	[53]
1986	$7+\epsilon$	Bombieri, Friedlander and Iwaniec	[15]
1996	6.9075	Fouvry and Grupp	[56]
1990	6.8354	Wu	[214]
1999	6.8325	Haugland	[94]

Now we will consider more general patterns. Let $\mathscr{H} := \{h_1, \ldots, h_k\}$ be a finite set of distinct nonnegative integers with $k \geq 2$. For each prime p, let $v_p(\mathscr{H})$ be the number of distinct residue classes modulo p of the elements of $\mathscr{H}$, and define a product