Twin progress in number theory

T.S. Trudgian*

There are many jokes of the form "X's are like buses: you wait ages for one and then n show up at once." There appear to be many admissible values of $\{X, n\}$: $\{\text{Ashes series}, 2\}$, $\{\text{efficiency dividends}, (m, \text{ where } m \to \infty)\}$.

Normally, when X is progress on a problem in number theory, n is a non-negative integer, strictly less than unity. Therefore it was with ebullience that, *on the same day*, I read of the proof of a centuries-old problem, and of admirable, and completely unexpected, progress towards a millenia-old problem. This short note attempts to explain the two problems and to give a brief outline of the methods used to tackle them.

Goldbach's Conjectures

The starting point with the primes is the following statement: every natural number $n > 1$ can be written as a product of primes in (essentially) exactly one way. It is necessary to throw in 'essentially', much in the same way as a lawyer throws in 'allegedly', to cover ourselves. If we disregard the order in which the prime factors of n appear then we can write n as a unique product of primes.[1]

Thus we can always *factorise* a number $n > 1$ into primes. Is it true that we can always *partition* a number $n > 1$ into primes? That is, can we write $n = p_1 + p_2 + \cdots$ for some primes $p_1, p_2, \ldots$? This is obviously true if we do not impose a limit on the number of primes: just take $p_1 = 2$ and, to obtain even numbers, add as many copies of 2 as you need; to obtain odd numbers, add copies of 3.

Instead, let k be a fixed number, possibly large, but finite. Can we partition every $n > 1$ as $n = p_1 + p_2 + \cdots + p_r$, where $r \leq k$? That is, can we write every integer greater than one as the sum of at most k primes? Clearly $k \geq 2$, since otherwise we could not partition composite numbers. Could it be that $k = 2$?

Let us examine this question according as n is even or odd. If n is even, and, were $k = 2$, then, apart from $n = 2$ and $n = 4$, we should require exactly two odd primes p_1 and p_2 that add to give n. This appears to stand up to some initial scrutiny, for example

$$6 = 3 + 3, \qquad 8 = 3 + 5, \qquad 10 = 3 + 7, \qquad 100 = 11 + 89.$$

*MSI, ANU College of Physical and Mathematical Sciences, Australian National University, Barry Drive, Acton, ACT 0200. Email: timothy.trudgian@anu.edu.au

[1]Incidentally, this is one reason why it is customary not to define 1 as a prime number.

That this ought always to be the case is encapsulated in

The Goldbach Conjecture. *Every even $n > 4$ is the sum of two odd primes.*

If n is odd and, were $k = 2$, then, apart from $n = 3$, we should require $n = 2 + p_1$, where p_1 is an odd prime. In particular we should require that $n - 2$ be prime for all n. That this statement is false infinitely often is apparent from considering $n \equiv 2 \pmod 3) = 2, 5, 8, \ldots$. Every second member of this sequence is odd, and yet $3 \mid (n - 2)$.

Therefore, when n is odd, we cannot hope for anything better than $k = 3$. Could it be that $k = 3$? If $n = p_1 + p_2 + p_3$, and n is odd, then either two of the primes are 2, or none of them is 2. Since the former case can be expurgated[2] we are left with the possibility that $n = p_1 + p_2 + p_3$, where all of the primes are odd. This appears to stand up to some initial scrutiny, for example

$$9 = 3 + 3 + 3, \qquad 11 = 3 + 3 + 5, \qquad 17 = 5 + 5 + 7, \qquad 101 = 3 + 19 + 79.$$

This brings us to

The Ternary Goldbach Conjecture. *Every odd $n > 7$ is the sum of three odd primes.*

The Goldbach Conjectures are named after Christian Goldbach who annunciated them in 1742. The Goldbach Conjecture implies the Ternary Goldbach Conjecture; for this reason the latter is sometimes called the Weak Goldbach Conjecture.[3]

The Goldbach Conjecture has been verified up to $n = 4 \times 10^{18}$ [5], and is widely believed to be true. Until May of this year, the Ternary Goldbach Conjecture was verified for $n \leq 8 \times 10^{26}$ [5], and also for $n \geq e^{3100} \approx 10^{1346}$ [4]. On 13 May, Harald Helfgott provided a proof [2] that filled the void. I should like to discuss briefly the method he used in his proof.

Small remark on technique: using the circle method

The starting point is to note that

$$\int_0^1 e^{2\pi i\alpha m}\,\mathrm{d}\alpha = \begin{cases} 1 & \text{if } m = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Given an $n > 1$, let $F(\alpha) = \sum_{p \leq n} e^{2\pi i\alpha p}$. It follows that $I(n) = \int_0^1 F(\alpha)^3 e^{-2\pi i\alpha n}\,d\alpha$ counts the number of ways that n can be written as a sum of three primes. We should like to show that, for all odd $n > 7$, $I(n) > 0$.

[2]Consider, for example, $n \equiv 4 \pmod 5$.

[3]Assume the Goldbach Conjecture, and let $n > 7$ be an odd integer. Then $n - 3$ is an even integer greater than 4, whence we have $n - 3 = p_1 + p_2$, where p_1 and p_2 are odd primes.

To accomplish this we partition the integral $I(n)$ into major and minor arcs. The major arcs consist of those real numbers α which we can approximate well by rational numbers with small denominators; the minor arcs are the remainder. To calculate the contribution along the major arcs we use a sum over the zeroes of a Dirichlet L-function. Helfgott's paper used extensive computations, performed with David Platt [3], over the zeroes of L-functions. To calculate the contribution along the minor arcs we need to use sieve methods. A sieve is a mathematical device used to select, from a large pool of numbers, certain numbers of a special shape.

Of course, none of this is easy to do: Helfgott describes his approach as a 'two-log gambit'. That is, he sacrifices two logarithms of accuracy early on in the piece in order to obtain greater flexibility and freedom of movement.

In 2012 Helfgott announced that he expected the proof to take a couple of years. I do not think that anyone expects progress on the Goldbach Conjecture in the next fifty years, let alone a proof of it.[4] However, this pessimism may well be unfounded. Indeed, I know of no one who expected progress to be made on the following problem. How wrong we all were!

Twin Primes

Consider integers n and $n+2$, where $n \geq 1$. If n is odd, then these numbers are co-prime. For, if there is a number d that divides both n and $n+2$, then d divides their difference, whence d is either 1 or 2: the latter may be ruled out since n is odd. Thus, for all $n \geq 1$, we have

$$n = \prod_i p_i^{e_i}, \qquad n+2 = \prod_j q_j^{f_j}, \tag{1}$$

where e_i and f_j are non-negative integers, and no p_i is a q_j. The simplest form that (1) may take is if $n = p$ and $n+2 = q = p+2$. Call such primes $p, p+2$ *twin primes.*

We can prove easily that there are infinitely many primes without a twin. It is known[5] that there are infinitely many primes of the form $p \equiv 1 \pmod 3$, whence infinitely often $p+2$ is divisible by 3.

On the other hand, there appears to be no shortage of twin primes: the largest known twin primes have over 200,000 digits. This brings us to a conjecture put forth by Euclid.

The Twin Prime Conjecture. *There are infinitely many twin primes.*

One could, with cavalier bravado, dare to make an even stronger conjecture. In 1849 Polignac conjectured that for all $k \geq 1$ the set $\{n, n+2k\}$ assumes prime

[4] Though I still get sent proofs from budding geniuses from time to time.

[5] The proof is quite quick provided one has the time to introduce quadratic reciprocity.

values infinitely often. That is, there are infinitely many twin primes $(p, p+2)$, infinitely many 'cousin primes' $(p, p+4)$, infinitely many sexy[6] primes $(p, p+6)$, etc.

Such problems seem out of reach at present. Since discretion is the better part of valour perhaps we should aim at the more modest

The Prime Gap Conjecture. *There exists some k for which there are infinitely many primes of the form $(p, p+a)$, where $a \leq k$.*[7]

Until recently it had been widely thought that, although this is clearly weaker than the Twin Prime Conjecture, it is still far too difficult to solve. In 2009, Goldston, Pintz and Yıldırım [1], under the assumption of a very powerful conjecture (the Elliot–Halberstam Conjecture[8]) proved that $k \leq 16$. There seemed no way to remove the reliance upon the Elliot–Halberstam Conjecture.

Seemingly out of nowhere, on 17 April, Yitang Zhang proved [6] that $k \leq 70,000,000$. His paper is full of dense, technical mathematics; the following section is an attempt to condense some of the ideas.

Primes in tuples

Let k_0 be a fixed number, possibly monstrously large, and let $\{h_i\}_{1 \leq i \leq k_0}$ be a set of non-negative integers. We wish to show that there are infinitely many n for which there are at least two primes in the set $\{n + h_i\}_{1 \leq i \leq k_0}$. Consider two sums

$$S_1 = \sum_{N < n \leq 2N} \lambda(n)^2, \qquad S_2 = \sum_{N < n \leq 2N} \left(\sum_{i=1}^{k_0} \theta(n + h_i) \right) \lambda(n)^2,$$

where the $\lambda(n)$ are real coefficients, referred to as *weights*, and where $\theta(n) = \log p$ if n is a prime, and zero otherwise. If one can show that

$$S_2 - (\log 3N) S_1 = \sum_{N < n \leq 2N} \lambda(n)^2 \left\{ \sum_{i=1}^{k_0} \theta(n + h_i) - \log 3N \right\} > 0, \qquad (2)$$

then there must be at least two primes in the set $\{n + h_i\}_{1 \leq i \leq k_0}$.[9]

The method of Goldston, Pintz, and Yıldırım looks at establishing (2) by careful estimates on the sums: this leads to

$$S_2 - (\log 3N) S_1 \geq M + O(E_1) + O(E_2),$$

[6] No giggling at the back: *sex* is Latin for 'six'.

[7] If you can prove this for $k = 2$, I shall happily co-write a paper with you.

[8] Which says, roughly speaking, that, when averaged over 'many' primes, primes in arithmetic progressions remain evenly distributed amongst all residue classes.

[9] There is at least one value of n for which $\sum_{i=1}^{k_0} \theta(n+h_i) - \log 3N$ is positive, whence there is at least one prime in the set $\{n + h_i\}$. Suppose that there is at most one prime contained therein, say $n + h_j$ is prime. Then $\theta(n + h_j) = \log(n + h_j) \leq \log(2N + h_j)$, and, for sufficiently large N, this is not greater than $\log 3N$ — a contradiction.

where M is a main term, and E_1 and E_2 are error terms. The aim of the game is to show that the main term is larger than both the error terms. A judicious choice of the weights $\lambda(n)$ is such that it involves $\log(D/d)$, where $d \leq D$, and where $D = N^{1/4+\epsilon_0}$, for some small, fixed, $\epsilon_0 > 0$. With this choice of D one can show that M is larger than E_1, but one cannot show that it is larger than E_2.

Zhang's idea was to add a restriction to the weight: not only should the sum be over $d \leq D$ but also over $d \mid \mathcal{P}$, where $\mathcal{P}$ is a product of the primes less than x^α for some small α. This was a classic gambit with an immediate loss and ultimate win.

Zhang now needed to prove that M exceeds E_1 *even with* this added restriction. Not only was Zhang able to do this, but the extra freedom wrought by this gambit enabled him to prove that M also exceeds E_2. Specifically, he was able to show that M dominates when $k_0 \geq 3,500,000$. From this he was able to deduce that there are infinitely many primes the distance between which is less than 70,000,000, or, equivalently, that

$$\liminf_{n\to\infty}(p_{n+1} - p_n) < 7 \times 10^7.$$

Zhang's result has prompted a lot of online activity in reducing the size of the prime gap. Much of this work is now part of the online collaboration *Polymath8*. To date, we know that there are infinitely many primes not more than $12,006$ apart.[10] You can follow these developments from the comfort of your own home: http://michaelnielsen.org/polymath1/index.php?title=Bounded_gaps_between_primes.

Closing remarks

Both Helfgott and Zhang introduce gambits into their gameplay. Such audacious tactics make for exciting mathematical arguments. Mathematical excitement is extremely important to focus the study of specialists on new aspects of a problem. Notwithstanding the healthy scepticism concomitant with the profession, unexpected progress can do wonders to shift us out of pessimism that progress will not be made in 'our lifetimes'. Bring on the twin primes!

References

[1] Goldston, D.A., Pintz, J. and Yıldırım, C.Y. (2009). Primes in tuples I. *Ann. Math.* **170,** 819–862.

[2] Helfgott, H. (2013). Major arcs for Goldbach's theorem. Available at http://arxiv.org/abs/1305.2897v2.

[3] Helfgott, H. and Platt, D.J. (2013). Numerical verification of the Ternary Goldbach Conjecture up to 8.875e30. Available at http://arxiv.org/abs/1305.3062.

[4] Liu, M.-Ch. and Wang, T.Z. (2002). On the Vinogradov bound in the three primes Goldbach conjecture. *Acta Arith.* **105,** 133–175.

[10] I should think it extremely unlikely that the current methods of optimising aspects of Zhang's paper would lead to a gap of size less than 1000.

[5] Silva, T.O.E., Herzog, S. and Pardi, S. (2013). Empirical verification of the Even Goldbach Conjecture, and computation of prime gaps, up to $4 \cdot 10^{18}$, 2013. *Math. Comp.* (to appear).
[6] Zhang, Y. Bounded gaps between primes. *Ann. Math.* (to appear).

Tim Trudgian is a number theorist who was an undergraduate at the ANU. He won a John Monash Scholarship to Oxford where he once took 6/24 in a friendly match in which the opposition may have reversed their batting order. After a brief stint in Canada, he returned to the ANU in 2012 as an ARC DECRA Fellow. He researches the Riemann zeta-function by day, and helps his son develop a sound cover-drive by night.