ADDITIVE PROBLEMS WITH PRIME VARIABLES
THE CIRCLE METHOD OF HARDY, RAMANUJAN AND LITTLEWOOD

ALESSANDRO ZACCAGNINI

ABSTRACT. In these lectures we give an overview of the circle method introduced by Hardy and Ramanujan at the beginning of the twentieth century, and developed by Hardy, Littlewood and Vinogradov, among others. We also try and explain the main difficulties in proving Goldbach's conjecture and we give a sketch of the proof of Vinogradov's three-prime Theorem.

1. ADDITIVE PROBLEMS

In the last few centuries many additive problems have come to the attention of mathematicians: famous examples are Waring's problem and Goldbach's conjecture. In general, an additive problem can be expressed in the following form: we are given $s \geq 2$ subsets of the set of natural numbers $\mathbb{N}$, not necessarily distinct, which we call $\mathcal{A}_1, \dots, \mathcal{A}_s$. We would like to determine the number of solutions of the equation

$$n = a_1 + a_2 + \cdots + a_s \tag{1.1}$$

for a given $n \in \mathbb{N}$, with the constraint that $a_j \in \mathcal{A}_j$ for $j = 1, \dots, s$, or, failing that, we would like to prove that the same equation has at least one solution for "sufficiently large" n. In fact, we can not expect, in general, that for very small n there will be a solution of equation (1.1). Furthermore, depending on the nature of the sets $\mathcal{A}_j$, there may be some arithmetical constraints on those n that may be "represented" in the form (1.1).

In Waring's problem we take an integer $k \geq 2$, and all sets $\mathcal{A}_j$ are equal to the set of the k-th powers of the natural numbers: the goal is to prove that there exists an integer $s(k)$ such that *every* natural number has a representation as a sum of at most s k-th powers. This has been proved by Hilbert by means of a very intricate combinatorial argument. Another interesting problem is the determination of the minimal value of s such that equation (1.1) has at least one solution for sufficiently large $n \in \mathbb{N}$, that is, allowing a finite set of exceptions. We recall Lagrange's four square theorem (every non negative integer can be written as the sum of four squares of non negative integers), and also that if we take $k = 2$ and $s = 2$, then the "arithmetical" set of exceptions contains the congruence class 3 mod 4.

In Goldbach's problem we set $\mathcal{A}_1 = \mathcal{A}_2 = \mathfrak{P}$, the set of all prime numbers, and, of course, we are interested only in even values of n in (1.1).

In both Waring and Goldbach's problems we may say, somewhat vaguely, that the difficulties arise from the fact that the sets $\mathcal{A}$ have a simple multiplicative structure, but we are *adding* their elements.

1.1. **The circle method.** The method that we are going to describe, that has been widely used to tackle and solve many additive problems, has its origin in a 1918 paper of Hardy & Ramanujan [23] on partitions. It has been developed by Hardy & Littlewood [21], [22] in the 1920's, and, because of their success, it is now referred to as the Hardy-Littlewood, or simply circle, method.

Date: March 28, 2012.

In what follows, we shall describe Hardy, Littlewood & Ramanujan's ideas in some detail. For the sake of simplicity, we begin with the case of a *binary* problem, that is, the case where $s = 2$. As a further simplification, we assume that $\mathcal{A}_1 = \mathcal{A}_2 = \mathcal{A}$. Of course, we also assume that $\mathcal{A}$ is an infinite set. We start by setting

$$f(z) = f_{\mathcal{A}}(z) \stackrel{\text{def}}{=} \sum_{n=0}^{+\infty} a(n) z^n, \qquad \text{where} \qquad a(n) = \begin{cases} 1 & \text{if } n \in \mathcal{A}, \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

Since $\mathcal{A}$ is infinite, the function f is a power series whose radius of convergence is 1 (it certainly has a singularity at $z = 1$, and it is regular for $|z| < 1$ by comparison with the sum of a geometric series). We are interested in the number of the representations of n in the form $a_1 + a_2$ with $a_j \in \mathcal{A}$ for $j = 1, 2$. Therefore, we set

$$r_2(n) \stackrel{\text{def}}{=} \bigl|\{(a_1, a_2) \in \mathcal{A} \times \mathcal{A} \colon n = a_1 + a_2\}\bigr|.$$

By the so-called Cauchy rule for the product of two absolutely convergent power series, when $|z| < 1$ we have

$$f^2(z) = \sum_{n=0}^{+\infty} c(n) z^n \qquad \text{where} \qquad c(n) = \sum_{\substack{0 \le h,k \le n \\ h+k=n}} \sum a(h) a(k). \tag{1.3}$$

Here $a(h)a(k) = 1$ if both h, $k \in \mathcal{A}$, and otherwise $a(h)a(k) = 0$, whence $c(n) = r_2(n)$. The same argument proves more generally that

$$f^s(z) = \sum_{n=0}^{+\infty} r_s(n) z^n \quad \text{where} \quad r_s(n) \stackrel{\text{def}}{=} \bigl|\{(a_1, \dots, a_s) \in \mathcal{A}^s \colon n = a_1 + \cdots + a_s\}\bigr|,$$

again for $|z| < 1$. By Cauchy's theorem, for $\rho < 1$ we have

$$r_2(n) = \frac{1}{2\pi i} \oint_{\gamma(\rho)} \frac{f^2(z)}{z^{n+1}} \, dz, \tag{1.4}$$

where $\gamma(\rho)$ is the circle whose centre is at the origin and whose radius is ρ. For some sets $\mathcal{A}$ it is possible to determine an asymptotic development for f around the singularities it has on the circle $\gamma(1)$, and it is therefore possible to estimate the integral in (1.4) taking ρ as a function of n whose limiting value is 1.

1.2. **A simple example.** As a simple example, we set up the circle method to solve a trivial combinatorial problem: given $k \in \mathbb{N}^*$, determine the number of possible representations of $n \in \mathbb{N}$ as a sum of exactly k natural numbers. In other words, we want to determine $r_k(n) := \bigl|\{(a_1, \dots, a_k) \in \mathbb{N}^k \colon n = a_1 + \cdots + a_k\}\bigr|$. It is clearly possible to show directly, in a completely elementary way, that this number is $r_k(n) = \binom{n+k-1}{k-1}$.

In this case we obviously have $f(z) = \sum_{n=0}^{+\infty} z^n = (1-z)^{-1}$, so that, for $\rho < 1$,

$$r_k(n) = \frac{1}{2\pi i} \oint_{\gamma(\rho)} \frac{dz}{(1-z)^k z^{n+1}}. \tag{1.5}$$

We remark that the integrand has only one singularity on the circle $\gamma(1)$, which is a pole. In this particular case it is possible to compute exactly the value of the integral on the right hand side of (1.5): in fact, since $\rho < 1$, we have the development

$$\frac{1}{(1-z)^k} = 1 + \binom{-k}{1}(-z) + \binom{-k}{2}(-z)^2 + \cdots = \sum_{m=0}^{+\infty} \binom{-k}{m} (-z)^m. \tag{1.6}$$

The series on the right converges uniformly on all compact sets contained in $\{z \in \mathbb{C}\colon |z| < 1\}$, and therefore we may substitute into (1.5) and interchange the integral and the series:

$$r_k(n) = \frac{1}{2\pi i} \sum_{m=0}^{+\infty} \binom{-k}{m} (-1)^m \oint_{\gamma(\rho)} z^{m-n-1} \, dz$$
$$= \frac{1}{2\pi i} \sum_{m=0}^{+\infty} (-1)^m \binom{-k}{m} \begin{cases} 2\pi i & \text{if } m = n, \\ 0 & \text{otherwise,} \end{cases} = (-1)^n \binom{-k}{n}.$$

It is not difficult to check that $(-1)^n \binom{-k}{n} = \binom{n+k-1}{k-1}$. We finally remark that the integrand is fairly small on the whole circle $\gamma(\rho)$, except for a small arc close to the point $z = \rho$, that gives the main contribution to the integral in (1.5). We will make things more precise later in (1.13).

In general, of course, it is not possible to evaluate directly and exactly the integral, and usually the integrand has several singularities on the circle $\gamma(1)$. For instance, in order to compute the number of possible decompositions of an integer $n \in \mathbb{N}$ as a sum of k odd integers, we need the function $g(z) = \sum_{m=0}^{+\infty} z^{2m+1} = z/(1-z^2)$, that has two singularities, namely $z = \pm 1$. In these cases, one needs asymptotic developments near each singularity. It is an interesting exercise to repeat the same computations as above in this case, to see how the arithmetical condition $n \equiv k \bmod 2$ arises.

We call *major arc* each region, close to a singularity, where a precise evaluation is performed. The complement of the set of major arcs is called *minor arcs*.

We notice a very important feature of (1.3) which we are going to exploit later when dealing with the Goldbach problem: when it is difficult to prove that $r(n) > 0$, it may be helpful to change the definition of the coefficients $a(h)$ in (1.2). Instead of allowing only the values 0 and 1, we may attach a positive weight to each element of the set $\mathcal{A}$: the resulting function will not count the number of representations anymore as $r_2(n)$ does, but it will be positive if and only if the weighted version is. The rationale is that it should be easier to bound from below a larger number, and, when dealing with primes it is well-known that θ is a simpler function than π. This gives the circle method some flexibility.

1.3. **Vinogradov's simplification.** The method we just roughly sketched has been used extensively by Hardy & Littlewood in the 1920's to prove many results connected to Waring's problem, and to carry out the first real attack on Goldbach's conjecture. In the 1930's Vinogradov introduced a few simplifications that make his method slightly simpler to explain. The basic idea in Hardy, Ramanujan and Littlewood is to have some fixed function, like $f(z)^k$ in the previous section, and to take ρ as a function of n with a limiting value of 1; furthermore, we need suitable asymptotic developments for f around the singularities that it has on the circle $\gamma(1)$. Vinogradov remarked that only the integers $m \leq n$ give a positive contribution to $r_2(n)$, as clearly shown by (1.3): following him, in the combinatorial problem of the previous section we introduce the function

$$f_N(z) \stackrel{\text{def}}{=} \sum_{m=0}^{N} z^m = \frac{1 - z^{N+1}}{1 - z}, \tag{1.7}$$

where the last equality is valid for $z \neq 1$. For $n \leq N$, Cauchy's theorem yields

$$r_k(n) = \frac{1}{2\pi i} \oint_{\gamma(1)} \frac{f_N^k(z)}{z^{n+1}} \, dz. \tag{1.8}$$

In this case there are *no* singularities of the integrand, since f_N is now a finite sum, a polynomial: therefore we may fix once and for all the circle of integration. Let us set $e(x) := e^{2\pi i x}$ and perform the change of variable $z = \mathrm{e}(\alpha)$ in (1.8):

$$r_k(n) = \int_0^1 f_N^k\big(\mathrm{e}(\alpha)\big)\,\mathrm{e}(-n\alpha)\,\mathrm{d}\alpha. \tag{1.9}$$

This is the Fourier coefficient formula, that gives the n-th coefficient in the Fourier series expansion of the periodic function $f_N^k\big(\mathrm{e}(\alpha)\big)$, because of the orthogonality property of the complex exponential function when n is an integer:

$$\int_0^1 \mathrm{e}(nx)\,\mathrm{d}x = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1.10}$$

Its importance lies in the fact that it transforms an arithmetical problem into one that can be attacked using standard techniques from real and complex analysis. For simplicity, we set $T_N(\alpha) = T(\alpha) := f_N\big(\mathrm{e}(\alpha)\big)$; from (1.7) we deduce that

$$\begin{aligned} T(\alpha) &\stackrel{\text{def}}{=} \sum_{m=0}^{N} \mathrm{e}(m\alpha) \\ &= \begin{cases} \dfrac{1-\mathrm{e}\big((N+1)\alpha\big)}{1-\mathrm{e}(\alpha)} = \mathrm{e}\big(\tfrac{1}{2}N\alpha\big)\dfrac{\sin(\pi(N+1)\alpha)}{\sin(\pi\alpha)} & \text{if } \alpha \notin \mathbb{Z}; \\ N+1 & \text{if } \alpha \in \mathbb{Z}. \end{cases} \end{aligned} \tag{1.11}$$

Figure 1 shows the graph of $|T_{20}(\alpha)|$. The property that we need to conclude our elementary analysis concerns the rate of decay of the function T as α gets away from integers: from (1.11) we easily get

$$|T_N(\alpha)| \le \min\left(N+1, \frac{1}{|\sin(\pi\alpha)|}\right) \le \min\big(N+1, \|\alpha\|^{-1}\big) \tag{1.12}$$

where $\|\alpha\|$ denotes the distance of α from the nearest integer, that is $\min\{\{\alpha\}, 1-\{\alpha\}\}$, since T is periodic of period 1, and $\alpha \le \sin(\pi\alpha)$ for $\alpha \in \big(0, \frac{1}{2}\big]$. This inequality shows that if $\delta = \delta(N)$ is not too small, the interval $[\delta, 1-\delta]$ does not give a large contribution to the integral in (1.9): in fact, if $\delta \ge 1/N$ and $k \ge 2$ we have

$$\left|\int_\delta^{1-\delta} T_N^k(\alpha)\mathrm{e}(-n\alpha)\,\mathrm{d}\alpha\right| \le \int_\delta^{1-\delta} |T_N^k(\alpha)|\,\mathrm{d}\alpha \le \int_\delta^{1-\delta} \frac{\mathrm{d}\alpha}{\|\alpha\|^k} \le \frac{2}{k-1}\delta^{1-k}, \tag{1.13}$$

and this is $o\big(N^{k-1}\big)$ as soon as $\delta^{-1} = o(N)$. In other words, it is sufficient that δ is just larger than N^{-1} so that the contribution of the interval $[\delta, 1-\delta]$ to the integral in (1.9) be smaller that the main term, that we know is $N^{k-1}(k-1)!^{-1}$. This means that the main term arises from a comparatively small interval around $\alpha = 0$.

In the case $k = 2$ we push our analysis a step farther: in fact, it is possible to prove (by induction plus some trigonometric identities) the formula

$$\left(\frac{\sin\big(\pi(N+1)\alpha\big)}{\sin(\pi\alpha)}\right)^2 = \sum_{|m| \le N+1} (N+1-|m|)\,\mathrm{e}(m\alpha). \tag{1.14}$$

Of course, the knowledge of this identity is at least as difficult as the knowledge of the correct answer to the original problem, as was the knowledge of (1.6). Indeed, a much more sensible approach would be to prove (1.14) by means of this argument, rather than the other way around. By (1.9) we have

$$r_2(n) = \int_0^1 \left(\frac{\sin\big(\pi(N+1)\alpha\big)}{\sin(\pi\alpha)}\right)^2 \mathrm{e}((N-n)\alpha)\,\mathrm{d}\alpha$$

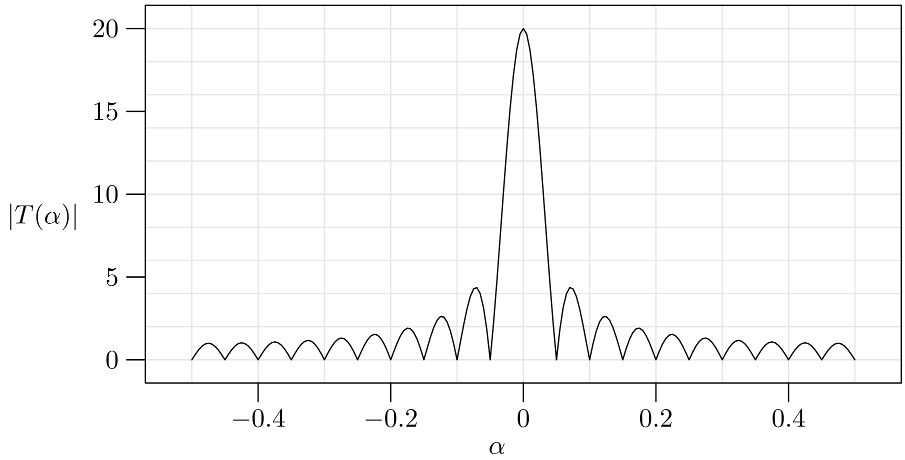

FIGURE 1. The graph of the function $|T_{20}(\alpha)|$, showing that it has a fairly large peak around integral values of α, and that otherwise it is comparatively small.

$$= \sum_{|m|\le N+1} (N+1-|m|) \int_0^1 \mathrm{e}((N+m-n)\alpha)\,\mathrm{d}\alpha.$$

By (1.10), the only non-vanishing integral occurs for $m = n-N$, so that $r_2(n) = N+1-|n-N| = n+1$. The point of this example is that one can usually find information on the quantity $r_2(n)$ by using transformations and suitable identities. We develop the subject further in the next section.

2. GOLDBACH'S PROBLEM

After this fairly long introduction devoted to the mechanism of the circle method, we now want to set it up in the case of the Goldbach's problem. Henceforward, the variables p, p_1, p_2, ..., always denote prime numbers. We are interested in the number of representations of n as a sum of two prime numbers

$$r_2(n) \stackrel{\mathrm{def}}{=} \left|\{(p_1,p_2) \in \mathfrak{P}\times\mathfrak{P} \colon n = p_1+p_2\}\right|,$$

where p_1 and p_2 are not necessarily distinct, but we consider p_1+p_2 and p_2+p_1 as distinct representations of n if $p_1 \ne p_2$. For the time being, we do not assume that n is an even integer. Goldbach's conjecture, as stated in a 1742 letter to Euler, is that $r_2(n) \ge 1$ for all even $n \ge 4$. A very crude argument, namely the consideration of the *average* number of representation of an even integer n as a sum of 2 primes, shows that we may reasonably expect that $r_2(n) \asymp n/(\log n)^2$. In fact, there are about $N/\log N$ odd primes p up to N, hence about $N^2/(\log N)^2$ couples (p_1,p_2) of odd primes both below N, and their sums are even numbers that do not exceed $2N$. We may therefore expect that, unless something weird happens, an even integer $n \le 2N$ should get its fair share of representations as p_1+p_2, that is, about $n/(\log n)^2$. This argument is quite crude, and neglects all *arithmetical* aspects of the problem, but it gives an idea of the expected size of $r_2(n)$, and, above all, confirms the expectation that not only $r_2(n) > 0$ for every even $n \ge 4$, but also that $r_2(n)$ tends to infinity quite rapidly, over the sequence of even numbers. In other words, a "typical" large even integer n should have $\asymp n/(\log n)^2$ Goldbach representations unless there is some obstruction, at present hidden, to prevent it.

For technical reasons that will be clarified later (essentially the same reason why it is easier to work with the Chebyshev θ or ψ functions rather than the π function) we prefer to consider

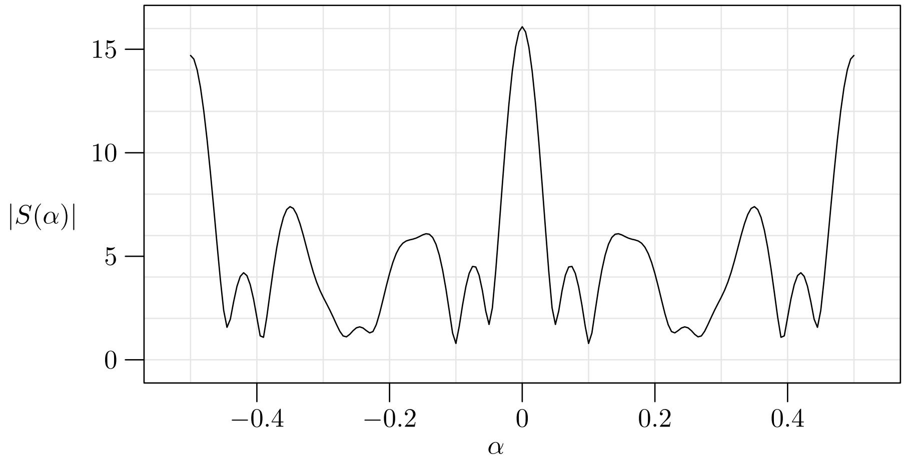

FIGURE 2. The graph of the function $|S_{20}(\alpha)|$ showing peaks close to the rational values $\alpha = 0, \frac{1}{2}, \pm\frac{1}{3}$; in $\alpha = \pm\frac{1}{4}$ there are no peaks because $\mu(4) = 0$.

a weighted version of the quantity $r_2(n)$, that is

$$R_2(n) \stackrel{\text{def}}{=} \sum_{p_1+p_2=n} \log p_1 \log p_2.$$

In other words, we count each representation of n as $p_1 + p_2$ with a weight $\log p_1 \log p_2$: this will make things easier, while still retaining the most important feature, that is, $r_2(n) > 0$ if and only if $R_2(n) > 0$. Therefore, the goal of the proof of Goldbach's conjecture in its original form, that $r_2(n) > 0$ for large even n, may be achieved by proving that $R_2(n) > 0$ for large even n. It is also clear that $r_2(n) \sim R_2(n)/(\log n)^2$ for large n. Using the traditional notation we set

$$S(\alpha) = S_N(\alpha) \stackrel{\text{def}}{=} \sum_{p \le N} (\log p)\, \mathrm{e}(p\alpha) \qquad \text{and} \qquad \theta(N;q,a) \stackrel{\text{def}}{=} \sum_{\substack{p \le N \\ p \equiv a \bmod q}} \log p.$$

By the orthogonality relation (1.10), for $n \le N$ we have

$$\int_0^1 S(\alpha)^2 \mathrm{e}(-n\alpha)\, \mathrm{d}\alpha = \sum_{p_1 \le N} \sum_{p_2 \le N} \log p_1 \log p_2 \int_0^1 \mathrm{e}\big((p_1 + p_2 - n)\alpha\big)\, \mathrm{d}\alpha = R_2(n). \tag{2.1}$$

Once again, this is the Fourier coefficient formula for the function S^2: compare (1.9). Since there are no singularities whatsoever on the circle of integration (though, strictly speaking, the circle has now been replaced by the interval $[0,1]$) we may wonder what plays the role of the major arcs: the distribution of prime numbers in arithmetic progressions now enters the picture, and we recall a basic result from analytic number theory.

Theorem 2.1. *For any fixed $A > 0$, there exists a constant $C = C(A) > 0$ such that for $N \to +\infty$ and uniformly for all $q \le (\log N)^A$ and for all integers a such that $(a,q) = 1$ we have*

$$\theta(N;q,a) = \frac{N}{\phi(q)} + E_1(N;q,a),$$

where

$$E_1(N;q,a) = O_A\Big(N \exp\{-C(A)\sqrt{\log N}\}\Big).$$

Before working it out in general, we compute $S(0)$, $S\big(\frac{1}{2}\big)$, $S\big(\frac{1}{3}\big)$, $S\big(\frac{1}{4}\big)$, and compare our result with the graph shown in Figure 2. It is quite straightforward that $S(0) = \theta(N;1,1) \sim N$

by Theorem 2.1; if we compute $S(\frac{1}{2})$ we see at once that

$$S\bigl(\tfrac{1}{2}\bigr) = \sum_{p\le N} (\log p)\,\mathrm{e}^{\mathrm{i}\pi p} = \log 2 - \sum_{\substack{p\le N\\ p\equiv 1 \bmod 2}} \log p = \log 2 - \theta(N;2,1) \sim -N,$$

since all primes $p \ge 3$ are odd so that $\mathrm{e}^{\mathrm{i}\pi p} = -1$. A quite similar thing happens if we compute $S(\frac{1}{3})$: exploiting the fact that $\mathrm{e}(x/3)$ has period 3, we see that

$$\begin{aligned} S\bigl(\tfrac{1}{3}\bigr) &= \log 3 + \sum_{\substack{p\le N\\ p\equiv 1 \bmod 3}} (\log p)\,\mathrm{e}^{2\mathrm{i}\pi p/3} + \sum_{\substack{p\le N\\ p\equiv 2 \bmod 3}} (\log p)\,\mathrm{e}^{2\mathrm{i}\pi p/3} \\ &= \log 3 + \mathrm{e}^{2\mathrm{i}\pi/3} \sum_{\substack{p\le N\\ p\equiv 1 \bmod 3}} \log p + \mathrm{e}^{4\mathrm{i}\pi/3} \sum_{\substack{p\le N\\ p\equiv 2 \bmod 3}} \log p \\ &= \mathrm{e}^{2\mathrm{i}\pi/3}\theta(N;3,1) + \mathrm{e}^{4\mathrm{i}\pi/3}\theta(N;3,2) + \log 3 \\ &= \bigl(\mathrm{e}^{2\mathrm{i}\pi/3} + \mathrm{e}^{4\mathrm{i}\pi/3}\bigr)\frac{N}{2} + O\Bigl(N\exp\{-C\sqrt{\log N}\}\Bigr) \qquad (2.2) \\ &= -\frac{N}{2} + O\Bigl(N\exp\{-C\sqrt{\log N}\}\Bigr), \end{aligned}$$

by Theorem 2.1. We leave the computation of $S(\frac{1}{4})$ as an exercise to the reader: the most important difference lies in the fact that the sum of roots of unity that occurs in (2.2) is replaced by $\mathrm{i} - \mathrm{i} = 0$, so that $S(\frac{1}{4}) = O\bigl(N\exp\{-C\sqrt{\log N}\}\bigr)$.

More generally, we now compute S at a rational number a/q, for $1 \le a \le q$ and $(a,q) = 1$:

$$\begin{aligned} S\Bigl(\frac{a}{q}\Bigr) &= \sum_{h=1}^{q} \sum_{\substack{p\le N\\ p\equiv h \bmod q}} (\log p)\,\mathrm{e}\bigl(p\tfrac{a}{q}\bigr) = \sum_{h=1}^{q} \mathrm{e}\bigl(h\tfrac{a}{q}\bigr) \sum_{\substack{p\le N\\ p\equiv h \bmod q}} \log p \\ &= \sum_{h=1}^{q} \mathrm{e}\bigl(h\tfrac{a}{q}\bigr)\theta(N;q,h) = \sum_{h=1}^{q}{}^{*}\, \mathrm{e}\bigl(h\tfrac{a}{q}\bigr)\theta(N;q,h) + O(\log q), \qquad (2.3) \end{aligned}$$

where the * means that we have attached the condition $(h,q) = 1$ to the corresponding sum. By Theorems 2.1 and A.1, and (2.3), for $q \le P := (\log N)^A$ we have

$$\begin{aligned} S\Bigl(\frac{a}{q}\Bigr) &= \frac{N}{\phi(q)} \sum_{h=1}^{q}{}^{*}\, \mathrm{e}\bigl(h\tfrac{a}{q}\bigr) + \sum_{h=1}^{q}{}^{*}\, \mathrm{e}\bigl(h\tfrac{a}{q}\bigr) E_1(N;q,h) + O(\log q) \\ &= \frac{\mu(q)}{\phi(q)} N + O\Bigl(NP\exp\{-C\sqrt{\log N}\}\Bigr), \qquad (2.4) \end{aligned}$$

where μ denotes the Möbius function. This formula suggests that $|S(\alpha)|$ is fairly large, unless of course $\mu(q) = 0$, when α is a rational number a/q, and that the size of $\bigl|S(a/q)\bigr|$ decreases essentially as q^{-1}. Since S is a continuous function, we may expect that $|S|$ be large in a neighbourhood of a/q, and we will exploit this fact to find an approximate formula for $R_2(n)$. We begin by extending the influence of the peak near a/q as much as possible: the simplest tool to use in this context is partial summation.

Lemma 2.2. *For any choice of $A > 0$, there exists a positive constant $C = C(A)$ such that for $1 \le a \le q \le P := (\log N)^A$, with $(a,q) = 1$ and for $|\eta| \le PN^{-1}$ we have*

$$S\Bigl(\frac{a}{q} + \eta\Bigr) = \frac{\mu(q)}{\phi(q)} T(\eta) + E_2(N;q,a,\eta) \qquad (2.5)$$

where

$$E_2(N;q,a,\eta) = O_A\Big(N\exp\{-C(A)\sqrt{\log N}\}\Big).$$

This is Lemma 3.1 of Vaughan [48]: the main ingredients for the proof are Theorem 2.1, partial summation, equation (2.4) and Theorem A.1. In a sense, the peak of S at a/q can be approximated fairly well by means of the peak of T at 0, after a suitable rescaling. Theorem 2.1 implies that the coefficient in $S(\alpha)$ is 1 on average, as the coefficient in $T(\alpha)$, which is a much easier function to work with.

For $q \le P$, we denote by $\mathfrak{M}(q,a) := \left[\frac{a}{q} - \frac{P}{N}, \frac{a}{q} + \frac{P}{N}\right]$ the *major arc* pertaining to the rational number with "small" denominator a/q, and write

$$\mathfrak{M} \stackrel{\text{def}}{=} \bigcup_{q\le P} \bigcup_{a=1}^{q}{}^{*} \mathfrak{M}(q,a) \qquad \text{and} \qquad \mathfrak{m} \stackrel{\text{def}}{=} \left[\frac{P}{N}, 1+\frac{P}{N}\right] \setminus \mathfrak{M},$$

where, once again, * means that we attach the condition $(a,q) = 1$. Therefore, $\mathfrak{M}$ is the set of the major arcs, and Lemma 2.2 suggests that it is the set where $|S|$ is fairly large and comparatively easy to approximate within reason. Its complement $\mathfrak{m}$ is the set of the *minor arcs*: we will see below in §3 that we do not have much control on the size of $|S|$ there. We translated the integration interval from $[0,1]$ to $\big[P/N, 1+P/N\big]$ in order to avoid two "half arcs" at 0 and 1, but this is legitimate since all functions involved have period 1.

The proof of this Lemma shows clearly that the major arcs can not be too numerous or too wide if we want to keep the resulting error term under control. We will use this result to find a quantitative version of Goldbach's conjecture, that was first justified along these lines by Hardy & Littlewood [21, 22].

For $n \le N$, from Equation (2.1) we deduce

$$\begin{aligned} R_2(n) &= \int_0^1 S(\alpha)^2 \mathrm{e}(-n\alpha)\,\mathrm{d}\alpha = \left(\int_{\mathfrak{M}} + \int_{\mathfrak{m}}\right) S(\alpha)^2 \mathrm{e}(-n\alpha)\,\mathrm{d}\alpha \\ &= \sum_{q\le P} \sum_{a=1}^{q}{}^{*} \int_{-P/N}^{P/N} S\left(\frac{a}{q}+\eta\right)^2 \mathrm{e}\Big(-n\big(\tfrac{a}{q}+\eta\big)\Big)\,\mathrm{d}\eta + \int_{\mathfrak{m}} S(\alpha)^2 \mathrm{e}(-n\alpha)\,\mathrm{d}\alpha \\ &= R_{\mathfrak{M}}(n) + R_{\mathfrak{m}}(n), \end{aligned}$$

say. From now on we write $\approx$ to indicate an expected asymptotic equality. For the time being we neglect the contribution of the minor arcs $R_{\mathfrak{m}}(n)$ and all the error terms that have arisen so far. By Equation (2.5) we have

$$\begin{aligned} R_{\mathfrak{M}}(n) &\approx \sum_{q\le P} \sum_{a=1}^{q}{}^{*} \int_{-P/N}^{P/N} \frac{\mu(q)^2}{\phi(q)^2} T(\eta)^2 \mathrm{e}\big(-n\big(\tfrac{a}{q}+\eta\big)\big)\,\mathrm{d}\eta \\ &= \sum_{q\le P} \frac{\mu(q)^2}{\phi(q)^2} \sum_{a=1}^{q}{}^{*} \mathrm{e}\big(-n\tfrac{a}{q}\big) \int_{-P/N}^{P/N} T(\eta)^2 \mathrm{e}(-n\eta)\,\mathrm{d}\eta. \end{aligned} \tag{2.6}$$

We extend the integral to the whole interval $[0,1]$ and recall the result from the previous section:

$$\int_0^1 T(\eta)^2 \mathrm{e}(-n\eta)\,\mathrm{d}\eta = \sum_{\substack{m_1+m_2=n \\ m_1\ge 0,\, m_2\ge 0}} 1 = n+1 \sim n. \tag{2.7}$$

Since $(P/N)\cdot N \to \infty$, we see that (1.13) implies that the interval $[P/N, 1-P/N]$ gives a contribution $o(n)$. Therefore we expect that

$$R_{\mathfrak{M}}(n) \approx n \sum_{q\le P} \frac{\mu(q)^2}{\phi(q)^2} \sum_{a=1}^{q}{}^{*} \mathrm{e}\big(-n\tfrac{a}{q}\big) = n \sum_{q\le P} \frac{\mu(q)^2}{\phi(q)^2} c_q(n), \tag{2.8}$$

where c_q is the Ramanujan sum defined in Theorem A.1. The next step is to extend the summation to all $q \geq 1$, with the idea of using Theorem A.2, since the summand is a multiplicative function of q by Theorem A.1: we skip the detailed proof that the error term arising from this operation is of lower order of magnitude.[1] Now, assuming that the error term can in fact be neglected, by Theorem A.2, the right hand side of Equation (2.8) becomes

$$\begin{aligned} R_{\mathfrak{M}}(n) &\approx n \sum_{q \leq P} \frac{\mu(q)^2}{\phi(q)^2} c_q(n) \approx n \sum_{q \geq 1} \frac{\mu(q)^2}{\phi(q)^2} c_q(n) \\ &= n \prod_p \bigl(1 + f_n(p) + f_n(p^2) + \dots\bigr) \end{aligned} \tag{2.9}$$

where the product is taken over all prime numbers and $f_n(q) := \mu(q)^2 c_q(n)/\phi(q)^2$. Obviously $f_n(p^\alpha) = 0$ for $\alpha \geq 2$, and for $\alpha = 1$ Theorem A.1 (or a direct computation) implies that

$$f_n(p) = \frac{\mu(p)^2}{\phi(p)} \frac{\mu\bigl(p/(p,n)\bigr)}{\phi\bigl(p/(p,n)\bigr)} = \begin{cases} \dfrac{1}{p-1} & \text{if } p \mid n, \\ -\dfrac{1}{(p-1)^2} & \text{if } p \nmid n. \end{cases}$$

If n is odd, the factor $1 + f_n(2)$ vanishes, and Equation (2.9) predicts that we should not expect any representation of n as a sum of two primes. In fact, if n is odd then $R_2(n) = 0$ if $n-2$ is not a prime number, and $R_2(n) = 2\log(n-2)$ if $n-2$ is a prime number: the result in (2.9) should be understood as $R_{\mathfrak{M}}(n) = o(n)$. Conversely, if n is even we may transform Equation (2.9) by means of some easy computation:

$$\begin{aligned} R_{\mathfrak{M}}(n) &\approx n \prod_{p \mid n} \Bigl(1 + \frac{1}{p-1}\Bigr) \prod_{p \nmid n} \Bigl(1 - \frac{1}{(p-1)^2}\Bigr) \\ &= 2n \prod_{\substack{p \mid n \\ p>2}} \Bigl(\frac{p}{p-1} \cdot \frac{(p-1)^2}{p(p-2)}\Bigr) \prod_{p>2} \Bigl(1 - \frac{1}{(p-1)^2}\Bigr) \\ &= 2C_0 n \prod_{\substack{p \mid n \\ p>2}} \frac{p-1}{p-2} = n\mathfrak{S}(n), \end{aligned} \tag{2.10}$$

where $2C_0$ is the so-called *twin-prime* constant, and $\mathfrak{S}(n)$ is the *singular series* defined by

$$C_0 \stackrel{\text{def}}{=} \prod_{p>2} \Bigl(1 - \frac{1}{(p-1)^2}\Bigr) \qquad \text{and} \qquad \mathfrak{S}(n) \stackrel{\text{def}}{=} 2C_0 \prod_{\substack{p \mid n \\ p>2}} \frac{p-1}{p-2}. \tag{2.11}$$

For n odd it is convenient to put $\mathfrak{S}(n) = 0$. Equation (2.10) is the asymptotic formula for $R_2(n)$ found by Hardy & Littlewood: of course, it would imply the truth of Goldbach's conjecture, but it is much stronger. In §3 we explain why, in the current state of knowledge, it is impossible to prove it. It is clear from (2.10) that the weighted number of representations depends on the size of n and also on its prime factorisation: it is a nice exercise in sieve theory to see why it has to be so.

We conclude this section noticing that Equation (2.1) implies that

$$R_2(n) \leq \int_0^1 |S(\alpha)|^2 \, \mathrm{d}\alpha = \sum_{p \leq N} (\log p)^2 \leq \theta(N) \log N \sim N \log N$$

[1] Actually, this is strictly true only on average over n: see Vaughan [48], Chapter 3.

by Theorem 2.1, so that for n close to N the expected asymptotic formula (2.10) does not differ too much from this upper bound. A bound of the right order of magnitude and with an explicit constant is provided by sieve methods: see Theorem 3.11 of Halberstam & Richert [19].

3. Overcoming the difficulties

For the sake of brevity, we only describe the two more important questions that remain to be settled: the approximation of the Chebyshev θ function, and the contribution of the minor arcs.

3.1. **Approximation of the Chebyshev theta function.** The approximation of θ provided by the Prime Number Theorem for Arithmetic Progressions 2.1 is quite weak for two main reasons: we remarked above that it is only valid in a fairly restricted range of values for q, and this forces a rather small choice of P, the parameter that we use to define the major arcs.

The second reason is that is that the upper bound known today for the error term is too large: in fact, it is conjectured that its true order of magnitude is much smaller. It is well known that the difference $\theta(N;q,a) - N/\phi(q)$ depends essentially on a sum whose summands have the shape $N^\rho/(\phi(q)\rho)$, where ρ denotes the generic complex zero of suitable Dirichlet L-functions. In the simplest case, when $q = a = 1$, the relation referred to can be written in the form

$$\theta(N) = N - \sum_{\substack{\rho\in\mathbb{C} \text{ s. t. } \zeta(\rho)=0 \\ \rho=\beta+i\gamma \\ |\gamma|\le T}} \frac{N^\rho}{\rho} + O\Big(\frac{N}{T}(\log N)^2 + \sqrt{N}\Big) \tag{3.1}$$

where $\rho = \beta + i\gamma$ is the generic zero of the Riemann zeta function with $\beta \in (0,1)$, and $T \le N$. This is known as the *explicit formula*, and it suggests that it might be a good idea to replace the function $T(\eta)$ defined in (1.11) with a different approximation for $S(\eta)$, namely

$$K(\eta) \stackrel{\text{def}}{=} \sum_{n\le N} \Big(1 - \sum_{|\gamma|\le T} n^{\rho-1}\Big) e(n\eta) \tag{3.2}$$

where the coefficient of $e(n\eta)$ is the derivative with respect to N of the first two terms in (3.1), evaluated at n, since if f is regular, then $\sum f(n) \sim \int f(t)\,dt$. The approximation for S given by (3.2) is valid only in a neighbourhood of 0, but we can find similar approximations valid on each major arc introducing the Dirichlet L-functions. Variants of this idea have been successfully used in several problems.

It is well known that the optimal distribution for prime numbers is achieved if *all* real parts β of all zeros $\rho = \beta + i\gamma$ of the ζ function with $\gamma \ne 0$ are equal to $\frac{1}{2}$ (Riemann Conjecture): in this case, $\theta(N) = N + O\big(N^{1/2}(\log N)^2\big)$. Similarly, if *all* zeros $\beta + i\gamma$ of all Dirichlet L-functions with $\beta \in (0,1)$ have real part $\frac{1}{2}$ (Generalised Riemann Conjecture), then for $q \le N$

$$\theta(N;q,a) = \frac{N}{\phi(q)} + O\Big(N^{1/2}(\log N)^2\Big). \tag{3.3}$$

The exponent of N in the error term of (3.3) is optimal, and it can not be replaced by a smaller one. In particular, Goldbach's Conjecture *does not* follow from the Generalised Riemann Conjecture (3.3). We finally remark that the general case $q > 1$ is harder than the case $q = 1$: in fact in the present state of knowledge it is still not possible to rule out the existence of a *real* zero $\beta \in (0,1)$ of some Dirichlet L-function, with β very close to 1. This, essentially, is the reason why we had to impose a rather severe limitation for q in Theorem 2.1. In fact, the contribution from this zero would be $\pm N^\beta/(\phi(q)\beta)$, that is, very close to the "main term" $N/\phi(q)$, and it might spoil the asymptotic formula for $\theta(N;q,a)$ for this particular value of q, with consequences on the asymptotic formula for $R_2(n)$. See the detailed discussion in §6.

3.2. **The contribution from minor arcs.** The main problem concerning minor arcs is that it is not possible to give an individual estimate for their contribution: it is comparatively easy to prove that *on average* over the integers $n \in [1,N]$ the minor arcs give a negligible contribution to $R_2(n)$, but it is not possible to prove the same thing for any single n. By the Fourier coefficient formula, Bessel's inequality and the Prime Number Theorem 2.1 with $q = 1$ we have

$$\sum_{n\le N}\Bigl|\int_{\mathfrak{m}} S(\alpha)^2 e(-n\alpha)\,d\alpha\Bigr|^2 \le \int_{\mathfrak{m}} |S(\alpha)|^4\,d\alpha \le \sup_{\alpha\in\mathfrak{m}} |S(\alpha)|^2 \int_0^1 |S(\alpha)|^2\,d\alpha$$
$$= O\Bigl(N\log N \sup_{\alpha\in\mathfrak{m}} |S(\alpha)|^2\Bigr).$$

Equation (2.4) suggests (and the following Lemma proves, albeit in a slightly weaker form) that the supremum in the last formula should be roughly N^2P^{-2}, since if $\alpha\in\mathfrak{m}$ then it is "close" to a rational with denominator $> P$.

Lemma 3.1 (Vaughan)**.** *For $1 \le a \le q \le N$ with $(a,q) = 1$ and $|\eta| \le q^{-2}$ we have*

$$S\Bigl(\frac{a}{q}+\eta\Bigr) \ll (\log N)^4\bigl(Nq^{-1/2}+N^{4/5}+N^{1/2}q^{1/2}\bigr).$$

This is Theorem 3.1 of Vaughan [48]. It implies that

$$\sum_{n\le N}|R_{\mathfrak{m}}(n)|^2 = \sum_{n\le N}\Bigl|\int_{\mathfrak{m}} S(\alpha)^2 e(-n\alpha)\,d\alpha\Bigr|^2 = O\bigl(N^3(\log N)^9P^{-1}\bigr), \qquad (3.4)$$

since every point on $[0,1]$ is within q^{-2} of a rational a/q (this is an elementary result of Dirichlet), and on the minor arcs $q > P$ by definition. In its turn, Equation (3.4) implies that for the majority of values $n \in [1,N]$ we have that $|R_{\mathfrak{m}}(n)|$ is of lower order of magnitude than the contribution from the major arcs provided by (2.9). A more precise statement follows in §4.

We remark that the measure of the minor arcs is $1+o(1)$, so that the major arcs represent only a tiny portion of the interval $[0,1]$.

4. Results for "almost all" even integers

In this section we find it convenient to incorporate "evenness" in the variable, and write $2n$ in place of n as earlier. The argument sketched in §2 is not strong enough to prove Goldbach's Conjecture, but it can still be used to prove some interesting, albeit weaker, results. In particular, we now prove that integers $2n$ such that $R_2(2n) = 0$ are comparatively rare: more precisely, let us set $\mathcal{E} := \{2n\colon R_2(2n) = 0\}$. This is the so-called "exceptional set" for Goldbach's problem. We will prove that, given $B > 0$, we have $|\mathcal{E}\cap[1,N]| = O_B\bigl(N(\log N)^{-B}\bigr)$.

Theorem 4.1. *Given $B > 0$ we have*

$$\sum_{2n\le N}\bigl|R_2(2n)-2n\mathfrak{S}(2n)\bigr|^2 \ll_B N^3\bigl(\log N\bigr)^{-B}. \qquad (4.1)$$

Notice that the trivial upper bound for the left-hand side of (4.1) is N^3. A bound $o\bigl(N^3\bigr)$ would yield an exceptional set of size $o(N)$.

Sketch of the proof. Using the ideas in §2, it is not too difficult to give a rigorous proof of the fact that for $2n \le N$ we have

$$R_{\mathfrak{M}}(2n) = 2n\mathfrak{S}(2n,P) + O_A\bigl(n(\log n)P^{-1}\bigr) \qquad (4.2)$$

using Lemma 2.2 and Equations (1.12), (2.6)–(2.7), where

$$\mathfrak{S}(m,P) \stackrel{\text{def}}{=} \sum_{q\le P}\frac{\mu(q)^2}{\phi(q)^2}c_q(m). \qquad (4.3)$$

Theorems A.2, A.1 and standard estimates concerning Euler's ϕ function, show that

$$\sum_{2n\le N} \bigl|\mathfrak{S}(2n,P)-\mathfrak{S}(2n)\bigr|^2 \ll N(\log N)^2P^{-1}. \tag{4.4}$$

The elementary inequality $|a+b+c|^2 \le 3\bigl(|a|^2+|b|^2+|c|^2\bigr)$ implies that

$$\begin{aligned}\sum_{2n\le N} \bigl|R_2(2n)-2n\mathfrak{S}(2n)\bigr|^2 &\ll \sum_{2n\le N} \bigl|R_{\mathfrak{M}}(2n)-2n\mathfrak{S}(2n,P)\bigr|^2 \\ &\qquad + \sum_{2n\le N} \bigl|2n\mathfrak{S}(2n,P)-2n\mathfrak{S}(2n)\bigr|^2 + \sum_{2n\le N} \bigl|R_{\mathfrak{m}}(2n)\bigr|^2 \\ &\ll N^3(\log N)^{2-2A} + N^3(\log N)^{2-A} + N^3(\log N)^{9-A} \\ &\ll N^3(\log N)^{9-A}\end{aligned}$$

by (3.4), (4.2)–(4.4). Theorem 4.1 follows choosing $A \ge B+9$. □

Finally, let $\mathcal{E}'(N) := \bigl\{2n \in \bigl[\frac12 N, N\bigr] \colon R_2(2n)=0\bigr\} = \mathcal{E}\cap\bigl[\frac12 N,N\bigr]$. Equation (2.11) implies that $\mathfrak{S}(2n) \ge 2C_0$, so that

$$\sum_{2n\le N} \bigl|R_2(2n)-2n\mathfrak{S}(2n)\bigr|^2 \ge \sum_{\substack{2n\le N\\ R_2(2n)=0}} |2C_0 2n|^2 \ge \sum_{\substack{N/2\le 2n\le N\\ R_2(2n)=0}} 8C_0^2n^2 \ge \frac12 C_0^2\bigl|\mathcal{E}'(N)\bigr|N^2,$$

and $|\mathcal{E}'(N)| = O_B\bigl(N(\log N)^{-B}\bigr)$ for any fixed $B>0$ by Theorem 4.1. The result for $\mathcal{E}\cap[1,N]$ follows by decomposing the interval $[1,N]$ into $O(\log N)$ intervals of type $\bigl[\frac12 M, M\bigr]$. It is easy to see that this argument proves rather more: in fact, let $\mathcal{E}''(N,X) := \bigl\{2n \in \bigl[\frac12 N,N\bigr] \colon |R_2(2n)-2n\mathfrak{S}(2n)| > X\bigr\}$ for some large X. Then, arguing as above, we see that

$$\sum_{2n\le N} \bigl|R_2(2n)-2n\mathfrak{S}(2n)\bigr|^2 \ge \sum_{2n\in\mathcal{E}''(N,X)} \bigl|R_2(2n)-2n\mathfrak{S}(2n)\bigr|^2 \ge X^2|\mathcal{E}''(N,X)|.$$

By Theorem 4.1 we have $|\mathcal{E}''(N,X)| \ll_B N^3X^{-2}(\log N)^{-B}$. Taking $X = N(\log N)^{-B/3}$, say, we deduce that $R_2(2n) = 2n\mathfrak{S}(2n) + O\bigl(N(\log N)^{-B/3}\bigr)$ for all even integers $2n\in[1,N]$ with *at most* $O\bigl(N(\log N)^{-B/3}\bigr)$ exceptions. In other words, the asymptotic formula holds, within a small error, for all even integers outside an exceptional set of small size. In a sense, the summation over n in (4.1) turns this into a *ternary* problem, like the one we are about to discuss.

5. Vinogradov's three-prime theorem

The circle method can be successfully applied to many different problems: for example, using a notation consistent with the one above, we have

$$R_3(n) \stackrel{\text{def}}{=} \sum_{p_1+p_2+p_3=n} \log p_1 \log p_2 \log p_3 = \int_0^1 S(\alpha)^3 e(-n\alpha)\,d\alpha$$

if $n\le N$. An argument similar to the one in the previous sections shows that $R_3(n)$ is well approximated by the contribution of the major arcs alone, and this yields

$$R_{3,\mathfrak{M}}(n) = \frac12 n^2\mathfrak{S}_3(n) + O_A\Bigl(n^2(\log n)^{-A}\Bigr), \tag{5.1}$$

for any positive A. Here $\mathfrak{S}_3(n)=0$ if n is even and

$$\mathfrak{S}_3(n) \stackrel{\text{def}}{=} \prod_{p\nmid n}\Bigl(1+\frac{1}{(p-1)^3}\Bigr)\cdot\prod_{p\mid n}\Bigl(1-\frac{1}{(p-1)^2}\Bigr)$$

if n is odd. Having three summands in place of two changes radically the nature of the problem: we are content to remark that in this case an *individual* upper bound for the contribution of the minor arcs to $R_3(n)$ is indeed possible. In fact, for $n \le N$ and $q > P$ Lemma 3.1 implies that

$$\Big| \int_{\mathfrak{m}} S(\alpha)^3 \mathrm{e}(-n\alpha)\,\mathrm{d}\alpha \Big| \le \sup_{\alpha\in\mathfrak{m}} |S(\alpha)| \int_0^1 |S(\alpha)|^2\,\mathrm{d}\alpha = O\Big(n^2(\log n)^4 P^{-1/2}\Big), \tag{5.2}$$

and this is of the same size of the error term in (5.1), provided that we take, as we may, P a large power of $\log N$.

Finally, we conclude noticing that a very simple computation shows that the twin-prime problem is naturally linked to Goldbach's conjecture: in fact, we immediately have

$$\theta_N(n) \stackrel{\text{def}}{=} \sum_{\substack{p_2\le N\\ p_2-p_1=n}} \log p_1 \log p_2 = \int_0^1 |S(\alpha)|^2 \mathrm{e}(-n\alpha)\,\mathrm{d}\alpha.$$

This means that the two problems are strictly related and are of the same degree of difficulty, at least in the quantitative form.

6. Stronger bounds for the exceptional set

The procedure explained in §4 allows one to obtain an exceptional set for the Goldbach problem of size $\ll_A NL^{-A}$ for any fixed $A > 0$: this is due to van der Corput [7], Estermann [12] and Tchudakoff [44] (late 1930's). It should be noticed that there is a link between the main parameters N and P and the size of the exceptional set, that turns out to be, essentially, N/P. In order to improve upon this, and to obtain an exceptional set of size $N^{1-\delta}$ for some $\delta > 0$, say, one should be able to take $P = N^{\delta}$. Reviewing the proof above, we see that the main obstacle to this is the level of uniformity in Theorem 2.1. We know that, at least unconditionally, this is the best that one can obtain (see the discussion in §20 of Davenport [8]). On the other hand, Vaughan's estimate in Lemma 3.1 gives exactly what is needed in general, provided that P is as large as a power of N.

A slightly stronger (in an appropriate sense) version of Theorem 2.1 can be obtained if one refrains from estimating the contribution of all zeros, as in (11) from §20 of Davenport, but is willing to compute exactly, as far as this is possible, the contribution of the so-called "exceptional" (or Siegel) zero. We now define precisely what this zero is.

Definition 6.1 (Exceptional zero)**.** *There is a positive constant c_1 such that there is at most one real primitive character $\widetilde{\chi}$ to a modulus $\widetilde{r} \le N$ such that $L(s,\widetilde{\chi})$ has a real zero $\widetilde{\beta}$ satisfying*

$$\widetilde{\beta} > 1 - \frac{c_1}{\log N}.$$

This is taken *verbatim* from Davenport, §14. It should be remarked that L-functions associated to real imprimitive character χ induced by $\widetilde{\chi}$ also vanish at $\widetilde{\beta}$. Of course, the conductor of χ is a multiple of $\widetilde{r}$.

Theorem 6.2. *For any fixed $C > 0$ there exists $D = D(C) > 0$ such that*

$$\theta(N;q,a) = \frac{N}{\phi(q)} - \frac{\widetilde{\chi}(a)N^{\widetilde{\beta}}}{\phi(q)\widetilde{\beta}} + O\Big(N\exp(-D(\log N)^{1/2})\Big)$$

uniformly for $q \le \exp(C(\log N)^{1/2})$ and $(a,q) = 1$. The term containing $\widetilde{\chi}$ is present only if there exists an exceptional zero for a modulus $\widetilde{r} \le N$ according to Definition 6.1, and $\widetilde{r} \mid q$.

This is (9) from §20 of Davenport [8], with ψ replaced by θ, which is legitimate in view of the size of the error term.

The idea of retaining one term arising from the zeros can be traced back to the partially successful attempt by Vaughan [47] to reduce the size of the exceptional set in Goldbach's problem to $N^{1-\delta}$. It was fully accomplished in the famous paper by Montgomery & Vaughan [39], and, anticipating a little, it has been brought to perfection by Pintz [41] in 2006. Of course, the treatment of this term entails some technical complications, and also a new, important tool.

Vaughan used Theorem 6.2 in order to obtain a much more precise approximation for S on the major arcs than the one we obtained in the previous sections. This entails a delicate analysis of the "secondary main term" that arises, as well as the singular series associated to it.

We notice that if n is not in the exceptional set $\mathcal{E}$ and the exceptional zero exists, then no asymptotic formula for $R_2(n)$ can be given, but only a positive lower bound for it, because the term containing the contribution of the exceptional zero turns out to be almost of the same size as the main term and it may affect its order of magnitude. See (6.2) below.

6.1. **Gallagher's Prime Number Theorem.** We do not describe Vaughan's original approach, but we just say a few words about the (technically more difficult) solution given by Montgomery & Vaughan in [39], where we will need Gallagher's Prime Number Theorem.

Lemma 6.3 (Gallagher)**.** *There exist positive absolute constants c_2 and c_3 such that*

$$\sum_{q\le P}\sum_{\chi(q)}{}^{*}\max_{x\le N}\max_{h\le N}\Bigl(h+\frac{N}{P}\Bigr)^{-1}\Bigl|\sum_{x-h}^{x}{}^{\#}\chi(p)\log p\Bigr| \ll G(P)\exp\Bigl\{-c_2\frac{\log N}{\log P}\Bigr\}, \qquad (6.1)$$

provided that $\exp\bigl(\sqrt{\log N}\bigr)\le P\le N^{c_3}$. *The symbol # attached to the sum means that when* $q=1$ *it must be replaced by*

$$\sum_{x-h}^{x}\log p-\sum_{\substack{x-h<n\le x\\ n>0}}1,$$

and when there is an exceptional character $\widetilde{\chi}$ *it must be replaced by*

$$\sum_{x-h}^{x}\widetilde{\chi}(p)\log p+\sum_{\substack{x-h<n\le x\\ n>0}}n^{\widetilde{\beta}-1}.$$

Finally,

$$G(T)=\begin{cases}1 & \text{if } \widetilde{\beta} \text{ does not exist,}\\ (1-\widetilde{\beta})\log T & \text{if } \widetilde{\beta} \text{ exists.}\end{cases}$$

This is Theorem 7 of Gallagher [15], in the modified form given by Montgomery & Vaughan (see Lemma 4.3 in [39]). With this modification, it becomes effective. We remark that the presence of the factor $G(P)$ in (6.1) is an occurrence of the so-called Deuring-Heilbronn phenomenon. If a "bad" exceptional zero exists, that is one which is very close to $s=1$, then the contribution of the other zeros to the left-hand side of (6.1) is "small." This is a crucial feature of the technique developed in [39].

We cut a very long (and technical) story short. Using Gallagher's Lemma 6.3 to bound error terms, Montgomery & Vaughan found that

$$R_{\mathfrak{M}}(n)=n\mathfrak{S}(n)+\widetilde{I}(n)\widetilde{\mathfrak{S}}(n)+\text{error term}, \qquad (6.2)$$

where $\widetilde{\mathfrak{S}}(n)$ is a *generalised singular series* satisfying $|\widetilde{\mathfrak{S}}(n)|\le\mathfrak{S}(n)$, and

$$\widetilde{I}(n)=\sum_{k\in(P,n-P)}\bigl(k(n-k)\bigr)^{\widetilde{\beta}-1}\le n^{\widetilde{\beta}}.$$

The summand $\widetilde{I}(n)\widetilde{\mathfrak{S}}(n)$ is present only if an exceptional zero (defined with respect to a power of the parameter P) occurs. In case of a "bad" exceptional zero $\widetilde{\beta}$, that is, if $\widetilde{\beta}$ is very close to 1, the two terms on the right of (6.2) might be of the same order of magnitude, and, as far as we can tell, $\widetilde{\mathfrak{S}}(n)$ could be negative. A careful analysis is needed to conclude that $R_2(n)$ is positive, although perhaps not of the "expected" magnitude $n\mathfrak{S}(n)$, for all even $n \le N$ outside an exceptional set $\mathcal{E}$ satisfying $|\mathcal{E} \cap [1,N]| \ll N^{1-\delta}$. The "localisation" of the exceptional zero (that is, a precise form of Definition 6.1) is a crucial feature of the proof. See §14 of [8].

7. The new explicit formula found by Pintz

Pintz recently found a new explicit formula for $R_2(n)$: it is Theorem 10 of [41]. The starting point is the same as in Vaughan and in Montgomery & Vaughan, namely the approximation of S by means of K, but instead of estimating the contribution of non-exceptional zeros, Pintz computes as exactly as possible the contribution of *each* zero, both of the zeta function and of suitable L-functions. This led him to a spectacular result for the exceptional set in Goldbach's problem, with the bound $|\mathcal{E} \cap [1,N]| \ll N^{2/3}$.

The explicit formula itself is rather cumbersome. Here we just show how to "guess" its shape, using the approximation $K(\eta)$ defined by (3.2), inserting it in the formula for the major arc around 0 and performing the integration disregarding error terms. This yields one of the terms in Pintz's explicit formula. Let $E(T) = \{\rho = \beta + i\gamma \colon \zeta(\rho) = 0,\ \beta \in [0,1],\ |\gamma| \le T\}$. Then, using (3.2), we have

$$\int_{-P/N}^{P/N} K^2(\eta) e(-n\eta)\,d\eta = \sum_{m_1 \le N} \sum_{m_2 \le N} \Bigl(1 - \sum_{\rho_1 \in E(T)} m_1^{\rho_1 - 1}\Bigr) \times \Bigl(1 - \sum_{\rho_2 \in E(T)} m_2^{\rho_2 - 1}\Bigr) \int_{-P/N}^{P/N} e((m_1 + m_2 - n)\eta)\,d\eta. \tag{7.1}$$

Following Pintz, for $\rho \in E_1(T) := \{1\} \cup E(T)$ we let

$$A(\rho) = \begin{cases} +1 & \text{if } \rho = 1, \\ -1 & \text{if } \rho \ne 1, \end{cases}$$

and we rewrite (7.1) as

$$\begin{aligned}
\int_{-P/N}^{P/N} K^2(\eta) e(-n\eta)\,d\eta &= \sum_{\substack{m_1 \le N \\ m_2 \le N}}\sum \ \sum_{\substack{\rho_1 \in E_1(T) \\ \rho_2 \in E_1(T)}}\sum A(\rho_1) m_1^{\rho_1 - 1} A(\rho_2) m_2^{\rho_2 - 1} \int_{-P/N}^{P/N} e((m_1 + m_2 - n)\eta)\,d\eta \\
&\approx \sum_{\substack{m_1 \le N \\ m_2 \le N}}\sum \ \sum_{\substack{\rho_1 \in E_1(T) \\ \rho_2 \in E_1(T)}}\sum A(\rho_1) m_1^{\rho_1 - 1} A(\rho_2) m_2^{\rho_2 - 1} \int_0^1 e((m_1 + m_2 - n)\eta)\,d\eta \\
&= \sum_{\rho_1, \rho_2 \in E_1(T)}\sum A(\rho_1) A(\rho_2) \sum_{\substack{m_1, m_2 \le N \\ m_1 + m_2 = n}}\sum m_1^{\rho_1 - 1} m_2^{\rho_2 - 1} \\
&\approx \sum_{\rho_1, \rho_2 \in E_1(T)}\sum A(\rho_1) A(\rho_2) \int_0^n t^{\rho_1 - 1} (n - t)^{\rho_2 - 1}\,dt \\
&= \sum_{\rho_1, \rho_2 \in E_1(T)}\sum A(\rho_1) A(\rho_2) n^{\rho_1 + \rho_2 - 1} \int_0^1 u^{\rho_1 - 1} (1 - u)^{\rho_2 - 1}\,du \\
&= \sum_{\rho_1, \rho_2 \in E_1(T)}\sum A(\rho_1) A(\rho_2) n^{\rho_1 + \rho_2 - 1} \mathrm{B}(\rho_1, \rho_2),
\end{aligned}$$

where B is the Euler Beta-function. This, apart from general lack of precision in estimating the ensuing error terms, is the expected contribution from the major arc around 0. A similar summand arises for each major arc, but with a weight of an arithmetical nature that gives rise to the so-called "generalised" singular series. Of course, the really deep question is how to keep the error term small. At the time of writing (March 2012) details have not yet been published.

The actual shape of Pintz's explicit formula is a linear combination of terms of the type $A(\rho_1)A(\rho_2)n^{\rho_1+\rho_2-1}\mathrm{B}(\rho_1,\rho_2)$ with coefficients given by suitable "generalised singular series" of arithmetical significance, as in (6.2). Here ρ_1 and ρ_2 range over the set of zeros (and poles) of Dirichlet L-functions associated to primitive characters whose conductor (essentially) divides n. The main term occurs for $\rho_1 = \rho_2 = 1$, that is, it is contributed by the pole of the Riemann zeta-function considered as a Dirichlet L-function associated to the primitive character modulo 1. As a corollary,

$$\left|R_{\mathfrak{M}}(n) - n\mathfrak{S}(n)\right| \le \varepsilon n\mathfrak{S}(n) \tag{7.2}$$

for all even $n \le N$, provided that $C(\varepsilon)n$ is not a multiple of all "generalised exceptional moduli" r_i for which a "generalised exceptional character" exist. Here $C(\varepsilon)$ is a suitable positive constant. We conclude observing that the number of such "generalised exceptional characters" depends only on ε, and their moduli are bounded from below. Another consequence of Pintz's explicit formula is that $R_{\mathfrak{M}}(n) \ll n\mathfrak{S}(n)$ for all even integers n. Pintz's formula with just the two summands with $\rho_1 = \rho_2 = 1$ and $\rho_1 = \rho_2 = \widetilde{\beta}$ reduces to (6.2).

A word of warning is in place, here. It is usually not too difficult to guess the shape of the main term in additive problems like Goldbach's, and also, as we have just seen, the shape of some "secondary" terms, although, as usual in Analytic Number Theory, the real difficulty lies in keeping the error terms under control: we recall what Hardy said (see pages 15–19 of [20]) concerning the comparative ease in guessing the deepest results in Analytic Number Theory.

7.1. **Approximations to the Goldbach Conjecture.** It may seem a bit absurd to speak about "approximations" to a problem where everything is discrete as in this case. In a sense, the ternary Goldbach problem discussed in §5 is an approximation, as Chen's result [6]: in the former case we have 3 variables instead of 2, in the latter we look for solutions of the equation $n = n_1 + n_2$ where the *total* number of prime factors of n_1 and n_2 is at most 3 instead of 2. The question is how to find something that is somehow "closer" to the binary Goldbach problem.

A somewhat loose way of comparing the difficulty of different additive problems is to introduce the concept of *density*. Recalling the notation in §1, we say that the set $\mathcal{A}_1$ has density α_1 if there are about N^{α_1} elements in $\mathcal{A}_1 \cap [1,N]$ (give or take some logarithmic factors, say). We associate to the equation (1.1) the density $\rho = \alpha_1 + \cdots + \alpha_s$, so that the binary Goldbach problem has density 2, the ternary Goldbach problem has density 3, and so on. As a rule of thumb, a large number of variables and a large density gives an easy additive problem, whereas a small number of variables combined with low density gives rise to a very difficult problem. Of course, the density has to be at least 1 if we want to be able to represent "almost all" integers, except for arithmetical constraints. In fact, the discussion in §1 and 2 shows that $n^{\rho-1}$ is, very roughly speaking, the *average* number of representations of n in the form (1.1). Of course, now we are neglecting arithmetical constraints.

Another type of approximation would be to consider the equation $n = p_1 + p_2 + m^k$ where k is a fixed, large integer. This is a ternary problem with density $\rho = 2 + 1/k$, which can be made as close to 2 as we please. However, Montgomery & Vaughan's technique already applies to the equation $n = p + m^k$ (see Brünner, Perelli & Pintz [5] and Vinogradov [49] for $k = 2$ and [50] for general $k \ge 3$) with some technical difficulties due to the low density $\rho = 1 + 1/k$, as well as to the fact that the truncated singular series associated to this problem is not absolutely convergent.

7.2. **The Goldbach-Linnik problem.** A different kind of approximation is given in Languasco, Pintz & Zaccagnini [29] on the so-called Goldbach-Linnik problem. The goal is to represent n as the sum of 2 primes and k powers of 2, so that the density is still $\rho = 2$, in the sense defined above. In this case we have proved that one can have both an asymptotic formula and an exceptional set of size $N^{3/5}(\log N)^{10}$. This is due to the fact that an arithmetical averaging allows us to take a larger value of the parameter P than Pintz could in his work on the exceptional set for the Goldbach problem in [41]. Summing up, although the density in the Goldbach-Linnik problem is the same as the density in the original Goldbach problem, the presence of even just one more summand (as in the case $k = 1$) drastically changes the result. In fact, the weighted number of representations of n as a sum of two powers of primes and a power of 2 is

$$R_1''(n) = \sum_{n_1+n_2+2^\nu=n} \Lambda(n_1)\Lambda(n_2) = \sum_{2^\nu \le n} R_2(n-2^\nu).$$

It is plain that the averaging process allows the presence of a small number of "exceptional" summands $R_2(n-2^\nu)$, that is summands for which the expected asymptotic formula (7.2) for the Goldbach problem fails, without affecting the main term for $R_1''(n)$. The event that $n-2^\nu$ is in the exceptional set for the Goldbach problem for "many" values ν should be, and in fact is, quite rare. That is why we get a much smaller exceptional set for the Goldbach-Linnik problem.

It is also important to notice that we do not directly use the properties of the exponential sum associated to the powers of 2, since its "major" and "minor" arcs are much more difficult to exploit than in the case of the standard Goldbach problem.

7.3. **The role of the Generalised Riemann Hypothesis.** If the Generalised Riemann Hypothesis (GRH) is true, then Hardy & Littlewood [21] have shown that it is possible to give a very strong bound for $|\mathcal{E} \cap [1,N]|$. Here we simply give the version of Vaughan's estimate in Lemma 3.1 which is valid under the assumption of GRH.

Lemma 7.1 (Goldston)**.** *Assume that GRH holds. Then, for $q \ge 1$ and $N \ge 2$ we have*

$$S\Bigl(\frac{a}{q}+\eta\Bigr) - \frac{\mu(q)}{\phi(q)} T(\eta) \ll q^{1/2} N^{1/2} \log(qN) \bigl(\log(qN) + N^{1/2}|\eta|^{1/2}\bigr).$$

This is Lemma 5 from Goldston [16], and it is a slight variant of Lemma 12 in Baker & Harman [1]. As a consequence, we can take more numerous and wider major arcs than in the unconditional case. Using Lemma 7.1 it can be shown that $|\mathcal{E} \cap [1,N]| \ll N^{1/2}(\log N)^2$. We refer to [16] for the details.

8. Reviving the old approach

The original setting of the circle method in the paper by Hardy & Ramanujan [23] and in the subsequent work by Hardy & Littlewood [21], [22], and Linnik [38] used infinite series, as we explained in §1. This approach is technically more difficult to handle, but it is also slightly more efficient and precise, especially when dealing with conditional problems, that is, when some assumption, such as the Riemann Hypothesis, is made. Linnik used this technique in [38] to give an alternative proof of the three-prime theorem of I. M. Vinogradov (see §5 above).

More recently, this approach has been used by A. I. Vinogradov in [49] when dealing with the exceptional set for $p+m^2$ (see above) without the excluded-zero technique. In the last few years, it has been used by Languasco & Zaccagnini in [36] and [37] to improve, albeit marginally, previous results by Bhowmik & Schlage-Puchta [2] and Friedlander & Goldston [13], respectively. We briefly describe one such result. Starting formally from the definition (see Goldston's review [17] of Fujii's paper [14]) one can guess that

$$\sum_{n\le N} R_2(n) = \frac{1}{2}N^2 - 2\sum_{\rho}\frac{N^{\rho+1}}{\rho(\rho+1)} + \text{an error term}$$

where ρ runs over the non-trivial zeros of the Riemann zeta-function. Assuming the Riemann Hypothesis, Fujii proved that the error term is $\ll N^{4/3}(\log N)^2$, and this was later improved by Bhowmik & Schlage-Puchta to $\ll N^{1/2}(\log N)^5$. Recently Languasco & Zaccagnini [36] obtained the bound $\ll N^{1/2}(\log N)^3$ which is essentially optimal in the present state of knowledge. The main reason for the improvement is the fact that one can avoid the use of Gallagher's Lemma (Lemma 1 of [15]) at the cost of using infinite series. Other similar formulae will appear in [32].

9. Diophantine problems

There is a variant of the circle method that can be used to study Diophantine inequalities: it has been introduced by Davenport & Heilbronn [9] in 1946. The integration on a circle, that can be equivalently performed on the interval $[0,1]$, is now replaced by integration on the whole real line. In this type of questions we can not count "exact hits" as in Goldbach or Waring's problems, but rather "near misses." Hence, we need a measure of "proximity" which can be provided in a number of ways. One could use the characteristic function $\chi_{[-\eta,\eta]}$ of the interval $[-\eta,\eta]$ as the tool for counting the number of solutions of the inequality $-\eta \le x \le \eta$. As we will see presently, though, one essential feature of this method is the rate of decay at infinity of the Fourier transform of this function, which is too slow. Hence, it will turn out to be more useful to introduce the function

$$\widehat{K}_\eta(\alpha) = \max\{0, \eta - |\alpha|\} \tag{9.1}$$

where $\eta > 0$. A simple computation shows that this is the Fourier transform of

$$K_\eta(\alpha) = \Bigl(\frac{\sin(\pi\alpha\eta)}{\pi\alpha}\Bigr)^2. \tag{9.2}$$

The crucial property, as we remarked above, is the rate of vanishing at infinity: we have

$$K_\eta(\alpha) \ll \min(\eta^2, |\alpha|^{-2}), \tag{9.3}$$

which is a trivial consequence of (9.2).

For brevity, we only consider problems with prime variables. We assume that we are given r non-zero real numbers $\lambda_1, \dots, \lambda_r$, and positive real numbers $k_1, \dots, k_r$. The goal is to approximate to a given real number γ by means of values of the form

$$F(p_1,\dots,p_k) = \lambda_1 p_1^{k_1} + \cdots + \lambda_r p_r^{k_r}, \tag{9.4}$$

where $p_1, \dots, p_r$ denote primes. Loosely speaking, there are two parameters that measure the relative difficulty of the problem. One, not surprisingly from what we saw in the classical case, is the number of variables r: the larger, the easier the problem. The other relevant quantity is what we called the "density" of the form F in (9.4), that is $\rho = \rho_F = 1/k_1 + 1/k_2 + \cdots + 1/k_r$, in agreement with the definition in §7.1. If ρ is "small" the goal is to show that

$$\bigl|\lambda_1 p_1^{k_1} + \lambda_2 p_2^{k_2} + \cdots + \lambda_r p_r^{k_r} - \gamma\bigr| < \eta \tag{9.5}$$

has infinitely many solutions for every fixed $\eta > 0$. If ρ is "large" one expects to be able to prove the stronger result that, in fact, some $\eta \to 0$ is admissible in (9.5), and, more precisely, it should be possible to take η as a small negative power of $\max_j p_j$.

Of course, some hypothesis on the irrationality of at least one ratio λ_i/λ_j is necessary, and also on signs, if one wants to approximate to all real numbers and not only some proper subset.

In fact, if all k_j were positive integers and all ratios λ_i/λ_j were rational, the values in (9.4) would all be multiple of some fixed, positive real number, and the approximation problem (9.5) could be insoluble for sufficiently small $\eta > 0$ and γ outside a discrete set. This hypothesis is usually exploited on the so-called "intermediate" arc.

One interesting feature of this method is that one shows explicitly (in a sense) a sequence of solutions of the inequality (9.5). In fact, while the results in the previous sections hold for all sufficiently large values of the main parameter N, in the present problem it will be enough to show that there is a solution of inequality (9.5) with, say, $p_j^{k_j} \in [\delta X_n, X_n]$ for all $j = 1, \dots, r$, where δ is a small positive constant, and X_n denotes a suitable sequence with limit $+\infty$. In practice, if we assume that, say, λ_1/λ_2 is irrational, we know that there exist infinitely many solutions of the inequality

$$\left|\frac{\lambda_1}{\lambda_2} - \frac{a}{q}\right| < \frac{1}{q^2},$$

with a and $q \in \mathbb{Z}$ and $(a,q) = 1$, and we take some fixed power of the values q, arranged in increasing order, as the sequence X (we drop the useless suffix n). We will adhere to tradition and consider X as the main parameter, and will define the other ones as suitable functions of X, although, from a logical point of view, we should use q.

9.1. **General setting.** Let $k \ge 1$ be an arbitrary real number. We need an exponential sum related to the primes, defined by

$$S_k(\alpha) = \sum_{\delta X \le p^k \le X} \log(p)\, \mathrm{e}(p^k \alpha) \tag{9.6}$$

where δ is a small, fixed positive constant, which may depend on the coefficients λ_j, and is needed for technical reasons. For any measurable set $\mathfrak{X} \subseteq \mathbb{R}$ we set

$$I(\eta, \mathfrak{X}) = I(\eta, \mathfrak{X}, \gamma) = \int_{\mathfrak{X}} S_{k_1}(\lambda_1 \alpha) \cdots S_{k_r}(\lambda_r \alpha) K_\eta(\alpha)\, \mathrm{e}(-\gamma\alpha)\, \mathrm{d}\alpha. \tag{9.7}$$

Using (9.2) and (9.6) we see that the number of solutions $\mathcal{N}(X)$ of the inequality (9.5) with $p_j^{k_j} \in [\delta X, X]$ for $j = 1, 2, \dots, r$, satisfies

$$\begin{aligned} I(\eta, \mathbb{R}) &= \sum_{\substack{\delta X \le p_j^{k_j} \le X \\ j=1,\dots,r}} \log(p_1) \cdots \log(p_r) \int_{\mathbb{R}} K_\eta(\alpha) \mathrm{e}\big((\lambda_1 p_1^{k_1} + \cdots + \lambda_r p_r^{k_r} - \gamma)\alpha\big)\, \mathrm{d}\alpha \\ &= \sum_{\substack{\delta X \le p_j^{k_j} \le X \\ j=1,\dots,r}} \log(p_1) \cdots \log(p_r) \max(0, \eta - |\lambda_1 p_1^{k_1} + \cdots + \lambda_r p_r^{k_r} - \gamma|) \\ &\le \eta (\log X)^r \mathcal{N}(X). \end{aligned}$$

Of course, in the last step we used the characterisation (9.1) for the Fourier transform of K. This means that a lower bound for $I(\eta, \mathbb{R})$ automatically provides a lower bound for $\mathcal{N}(X)$. We will show below that it is possible to prove a lower bound of the expected order of magnitude for a sequence of values of X that depends on the continued-fraction expansion for λ_1/λ_2. This will yield infinitely many solutions to the inequality (9.5).

We decompose the real line into three parts: a neighbourhood of the origin, which is called "major arc" $\mathfrak{M}$ and that provides the main term for I; An "intermediate arc" $\mathfrak{m}$ consisting of an interval on each side of the major arc: the contribution of this set turns out to be small if X belongs to the sequence of values referred to above; Finally, a "trivial arc" $\mathfrak{t}$ which contributes very little, and where we use the upper bound (9.3) for $\widehat{K}$ defined in (9.1).

9.2. **The major arc.** We will approximate to the exponential sum S_k defined in (9.6) using the corresponding exponential sum, with the coefficient $\log p$ replaced by its average 1, as in the previous sections, and the relevant exponential integral:

$$U_k(\alpha) = \sum_{\delta X \le n^k \le X} \mathrm{e}(n^k\alpha) \qquad \text{and} \qquad T_k(\alpha) = \int_{(\delta X^{1/k})}^{X^{1/k}} \mathrm{e}(t^k\alpha)\,\mathrm{d}t.$$

Gallagher's Lemma (Lemma 1 of [15]) allows us to connect the mean-square average of $S_k - U_k$ to the the so-called Selberg integral, instead of using a pointwise bound: see Brüdern, Cook & Perelli [4]. Here we just describe the standard case $k = 1$, although more general cases have appeared: see Appendix B for a list of papers. The difference between T_k and U_k is very small by Euler's summation formula and we will forget about it altogether. For $h \le X$ let

$$J(X,h) \stackrel{\mathrm{def}}{=} \int_X^{2X} |\theta(x+h) - \theta(x) - h|^2 \,\mathrm{d}x.$$

In view of the Prime Number Theorem, one may expect that $J(X,h) = o\big(Xh^2\big)$, at least if h is not too small with respect to X. For $h \ge X^{\varepsilon}$, the bound $J(X,h) \ll_{\varepsilon} Xh^2$ follows immediately from the Brun-Titchmarsh inequality (see Theorem 3.7 in Halberstam & Richert [19]). The strongest known bound for J is given in Saffari & Vaughan [43], although the author [52] has given a weaker result, but still with $J(X,h) = o\big(Xh^2\big)$, in a slightly wider range for h.

The precise definition of the major arc $\mathfrak{M}$ depends on the uniformity in h of the bounds for $J(X,h)$ referred to above. Its width, that is crucial for the strength of the final result, can be chosen larger if an average result like the bound for the Selberg integral is used. In practice one may replace each S_k in $I(\eta,\mathfrak{M})$ as defined in (9.7) by the corresponding T_k: the ensuing integral can be extended to the whole real line with a small error term (thanks again to (9.3)) and it is fairly easy to compute exactly with a little care: it turns out to be $\gg \eta^2 X^{\rho-1}$, where ρ denotes the quantity that we called density above.

9.3. **The intermediate arc.** Assume for simplicity that $k_1 = k_2 = 1$. Here we exploit the hypothesis on the irrationality of λ_1/λ_2. It is a comparatively easy deduction from Vaughan's Lemma 3.1 that $|S_1(\lambda_1\alpha)|$ and $|S_1(\lambda_2\alpha)|$ can not both be large if $\alpha \in \mathfrak{m}$. The reason why this has to be true is that if $\alpha \in \mathfrak{m}$ and we have integers a_1, a_2, q_1 and q_2 such that

$$\big|\lambda_i\alpha q_i - a_i\big| < \frac{Q}{X}, \qquad \text{for } i = 1 \text{ and } 2,$$

and Q is chosen carefully, then at least one between q_1 and q_2 must be larger than Q. Vaughan's Lemma 3.1 then implies that the corresponding $S_1(\lambda_i\alpha)$ can not be too large. That suitable integers a_i and q_i exist follows from Dirichlet's approximation Theorem.

9.4. **The trivial arc.** Here the rate of decay of $\widehat{K}$ plays an essential role. We recall that the Fourier transform of the characteristic function of the interval $[-\eta,\eta]$, which would be the natural function to use in place of K_η, decays too slowly. There is a connection between regularity in Sobolev spaces and rate of decay at infinity of Fourier transforms. A rather general result of this type is Lemma 1 of Tolev [46].

Appendix A. Some useful results

Theorem A.1 (Ramanujan)**.** *The Ramanujan sum $c_q(n)$ defined below is a multiplicative function of q, and*

$$c_q(n) \stackrel{\mathrm{def}}{=} \sum_{h=1}^{q}{}^{*}\, \mathrm{e}\Big(\frac{hn}{q}\Big) = \mu\Big(\frac{q}{(q,n)}\Big) \frac{\phi(q)}{\phi\big(q/(q,n)\big)}.$$

Theorem A.2 (Euler Product)**.** *Let f be a multiplicative function, and assume that the series $\sum_{n\ge 1} f(n)$ is absolutely convergent. Then the following identity holds*

$$\sum_{n\ge 1} f(n) = \prod_p \Bigl(1 + f(p) + f(p^2) + f(p^3) + \cdots\Bigr),$$

where the product is taken over all prime numbers and is absolutely convergent.

Appendix B. Further reading

The standard reference for the circle method is Vaughan's monograph [48]: see in particular Chapter 1. See also Hardy [20] Chapter 8 (in particular §§8.1–8.7), James [26] §5, and Ellison [11] for the history of Waring's problem. The genesis of the idea of studying the behaviour of the generating function in the neighbourhood of many singularities is clearly explained in Hardy & Ramanujan [23] (in particular §§1.2–1.5) and in Hardy [20] Chapter 8 (in particular §§8.6–8.7). For Waring's problem see Hardy & Wright [24] Chapters 20–21 for an introduction, and Vaughan [48] for a detailed study. For the relationship between Laurent series and Fourier series see Titchmarsh [45] §13.12. See also the survey by Kumchev & Tolev [27].

Set $\mathcal{E} := \{2n \colon r_2(2n) = 0\}$. The complete detailed proof that for any $A > 0$ we have $|\mathcal{E} \cap [1,N]| = O_A\bigl(N(\log N)^{-A}\bigr)$ is in §3.2 of Vaughan [48]. Montgomery & Vaughan [39] proved the stronger result that $|\mathcal{E} \cap [1,N]| \ll N^{1-\delta}$ for some $\delta > 0$. A discussion of many problems related to variants of the Goldbach Conjecture can be found in Languasco [28], while Zaccagnini [51] deals with "mixed" problems with primes and powers. A heuristic argument in favour of the twin-prime conjecture can be found in Hardy & Wright [24], §22.20. See the introduction of Halberstam & Richert [19] for the general setting of the Schinzel & Sierpiński's conjectures and the notes for a quantitative version of the same conjectures due to Bateman & Horn. An upper bound for $r_2(n)$ of the correct order of magnitude is contained in Theorem 3.11. See Zaccagnini [53] for an elementary heuristic argument (based on a variant of Eratosthenes' sieve) in support of the asymptotic formula (2.10): in particular see Equations (6), (8) and (10), and the "Coda." Other strategies for the proof of Goldbach's Conjecture are discussed in Ribenboim [42] §4.VI. See also Guy [18] §C.1 for further references.

For Equations (3.3) and (3.4) see Davenport [8] Chapter 20 and Chapter 25 respectively. Chen [6] proved that every large even integer can be written as a sum of a prime and of an integer with at most 2 prime factors: see Halberstam & Richert [19] Chapter 10, or Bombieri [3] §9 for a comparatively simple proof with 4 in place of 2.

Equation (5.2) is in Davenport [8] Chapter 26. The Ternary Goldbach Problem is discussed in [8] Chapter 26 or [48] §3.1. Deshouillers, Effinger, te Riele & Zinoviev [10] proved that if the Generalised Riemann Conjecture is true then *every* odd integer $n \ge 7$ is a sum of three primes. Theorem A.1 is Theorem 272 of Hardy & Wright [24]. For useful results on the distribution of primes or the properties of the Riemann zeta function, see [8] Chapters 7–18, or Ivić [25] Chapters 11–12.

The the excluded-zero technique, introduced by Vaughan in [47], has been exploited also in Montgomery & Vaughan [39], and developed in Brünner, Perelli & Pintz [5]. It has also been used in [50], [54], and in Perelli & Zaccagnini [40]. In several of these papers, though, the set of excluded zeros is large, and we need powerful density theorems for zeros of Dirichlet L-functions to estimate their total contribution.

For the Davenport-Heilbronn variant of the circle method see Vaughan [48], Chapter 11. For other recent results concerning Diophantine problems with prime variables see Languasco & Settimi [30], Languasco & Zaccagnini [31], [33], [35], [34]. The Introduction of [34] contains full references to recent work in the field.

Acknowledgements: These notes were written while preparing lectures given at the "Harish-Chandra Research Institute," Allahabad, in 2005, and enlarged for another series of lectures given at the "Institute for Mathematical Sciences," Chennai, in 2012. The Author thanks both Institutions for their hospitality and excellent working environment, and also the "Istituto Nazionale di Alta Matematica" (INdAM) in Rome for funding one of the travels to India.

REFERENCES

[1] R. C. Baker and G. Harman, *Diophantine approximation by prime numbers*, J. London Math. Soc. (2) **25** (1982), 201–215.
[2] G. Bhowmik and J.-C. Schlage-Puchta, *Mean representation number of integers as the sum of primes*, Nagoya Math. J. **200** (2010), 27–33.
[3] E. Bombieri, *Le Grand Crible dans la Théorie Analytique des Nombres*, Societé Mathématique de France, Paris, 1974, Astérisque n. 18.
[4] J. Brüdern, R. J. Cook, and A. Perelli, *The values of binary linear forms at prime arguments*, Proc. of Sieve Methods, Exponential sums and their Application in Number Theory (G. R. H. Greaves *et al*, ed.), Cambridge University Press, 1997, pp. 87–100.
[5] R. Brünner, A. Perelli, and J. Pintz, *The exceptional set for the sum of a prime and a square*, Acta Math. Hungar. **53** (1989), 347–365.
[6] J. R. Chen, *On the representation of a larger even integer as the sum of a prime and the product of at most two primes*, Sci. Sinica **16** (1973), 157–176.
[7] J. G. van der Corput, *Sur l'hypothèse de Goldbach pour presque tous les nombres pairs*, Acta Arith. **2** (1937), 266–290.
[8] H. Davenport, *Multiplicative Number Theory*, third ed., Graduate Texts in Mathematics, vol. 74, Springer, 2000.
[9] H. Davenport and H. Heilbronn, *On indefinite quadratic forms in five variables*, J. London Math. Soc. **21** (1946), 185–193.
[10] J.-M. Deshouillers, G. Effinger, H. te Riele, and D. Zinoviev, *A complete Vinogradov 3-primes Theorem under the Riemann Hypothesis*, Electr. Res. Announcements American Mathematical Society **3** (1997), 99–104.
[11] W. J. Ellison, *Waring's problem*, Amer. Math. Monthly **78** (1971), 10–36.
[12] T. Estermann, *On Goldbach's problem: Proof that almost all even integers are sums of two primes*, Proc. London Math. Soc. (2) **44** (1938), 307–314.
[13] J. B. Friedlander and D. A. Goldston, *Sums of three or more primes*, Trans. Amer. Math. Soc. **349** (1997), no. 1, 287–310.
[14] A. Fujii, *An additive problem of prime numbers*, Acta Arith. **58** (1991), no. 2, 173–179.
[15] P. X. Gallagher, *A large sieve density estimate near* $\sigma = 1$, Invent. Math. **11** (1970), 329–339.
[16] D. A. Goldston, *On Hardy and Littlewood's contribution to Goldbach's conjecture*, Proceedings of the Amalfi Conference on Analytic Number Theory, held at Maiori, Amalfi, Italy, from 25 to 29 September, 1989 (E. Bombieri, A. Perelli, S. Salerno, and U. Zannier, eds.), Università di Salerno, 1992, pp. 115–155.
[17] D. A. Goldston, *Review of A. Fujii's paper* [14], 1992, MR1121079 (92k:11106).
[18] R. K. Guy, *Unsolved Problems in Number Theory*, second ed., Springer, 1994.
[19] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.
[20] G. H. Hardy, *Ramanujan. Twelve lectures on subjects suggested by his life and works*, third ed., Chelsea, New York, 1999.
[21] G. H. Hardy and J. E. Littlewood, *Some problems in "Partitio Numerorum"; III. On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.
[22] G. H. Hardy and J. E. Littlewood, *Some problems of "Partitio Numerorum"; V. A further contribution to the study of Goldbach's problem*, Proc. London Math. Soc. (2) **22** (1923), 46–56.
[23] G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatory analysis*, Proc. London Math. Soc. (2) **17** (1918), 75–115, = S. Ramanujan, "Collected papers," edited by G. H. Hardy, P. V. Seshu Aiyar and B. M. Wilson, Third ed., AMS–Chelsea, 1999; n. 36, 276–309.
[24] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, fifth ed., Oxford Science Publications, Oxford, 1979.
[25] A. Ivić, *The Theory of the Riemann Zeta–Function*, J. Wiley, New York, 1985.
[26] R. D. James, *Recent progress in the Goldbach problem*, Bull. Amer. Math. Soc. **55** (1949), 246–260.

[27] A. V. Kumchev and D. I. Tolev, *An invitation to additive prime number theory*, 2004, See http://www.ma.utexas.edu/~kumchev/A19.ps.
[28] A. Languasco, *Some results on Goldbach's problem*, Rend. Sem. Mat. Univ. Pol. Torino **53 (4)** (1995), 325–337.
[29] A. Languasco, J. Pintz, and A. Zaccagnini, *On the sum of two primes and k powers of two*, Bull. London Math. Soc. **39** (2007), 771–780.
[30] A. Languasco and V. Settimi, *On a diophantine problem with one prime, two squares of primes and s powers of two*, Acta Arith. (2012), ?–?
[31] A. Languasco and A. Zaccagnini, *On a diophantine problem with two primes and s powers of* 2, Acta Arith. **145** (2010), 193–208.
[32] A. Languasco and A. Zaccagnini, *A Cesàro average of Goldbach numbers*, In preparation, 2012.
[33] A. Languasco and A. Zaccagnini, *A Diophantine problem with a prime and three squares of primes*, Submitted for publication, 2012.
[34] A. Languasco and A. Zaccagnini, *A Diophantine problem with prime variables*, Submitted for publication, 2012.
[35] A. Languasco and A. Zaccagnini, *L^2-norms of exponential sums over prime powers*, Submitted for publication, 2012.
[36] A. Languasco and A. Zaccagnini, *The number of Goldbach representations of an integer*, Proc. Amer. Math. Soc. **140** (2012), 795–804.
[37] A. Languasco and A. Zaccagnini, *Sums of many primes*, J. Number Theory **132** (2012), 1265–1283.
[38] Yu. V. Linnik, *A new proof of the Goldbach-Vinogradow theorem*, Rec. Math. [Mat. Sbornik] N.S. **19** (1946), 3–8.
[39] H. L. Montgomery and R. C. Vaughan, *The exceptional set in Goldbach's problem*, Acta Arith. **27** (1975), 353–370.
[40] A. Perelli and A. Zaccagnini, *On the sum of a prime and a k–th power*, Izv. Ross. Akad. Nauk, Ser. Math. **59** (1995), 185–200.
[41] J. Pintz, *Recent results on the Goldbach conjecture*, Proceedings of the ELAZ Conference, May 24–28, 2004 (Stuttgart) (W. Schwarz and J. Stending, eds.), Franz Steiner Verlag, 2006, pp. 220–254.
[42] P. Ribenboim, *The New Book of Prime Numbers Records*, Springer, New York, 1996.
[43] B. Saffari and R. C. Vaughan, *On the fractional parts of x/n and related sequences. II*, Ann. Inst. Fourier **27** (1977), 1–30.
[44] N. Tchudakoff, *On the density of the set of even numbers which are not representable as a sum of two odd primes*, Izv. Akad. Nauk SSSR. Ser. Mat. **2** (1938), 25–40.
[45] E. C. Titchmarsh, *The Theory of Functions*, second ed., Oxford University Press, Oxford, 1988.
[46] D. I. Tolev, *On a diophantine inequality involving prime numbers*, Acta Arith. **51** (1992), 289–306.
[47] R. C. Vaughan, *On Goldbach's problem*, Acta Arith. **22** (1972), 21–48.
[48] R. C. Vaughan, *The Hardy–Littlewood Method*, second ed., Cambridge University Press, Cambridge, 1997.
[49] A. I. Vinogradov, *On a binary problem of Hardy and Littlewood*, Acta Arith. **46** (1985), 33–56 (Russian).
[50] A. Zaccagnini, *On the exceptional set for the sum of a prime and a k-th power*, Mathematika **39** (1992), 400–421.
[51] A. Zaccagnini, *Additive problems with prime numbers*, Rend. Sem. Mat. Univ. Pol. Torino **53** (1995), 471–486, Atti del "Primo Incontro Italiano di Teoria dei Numeri," Roma, 3–5 gennaio 1995.
[52] A. Zaccagnini, *Primes in almost all short intervals*, Acta Arith. **84.3** (1998), 225–244.
[53] A. Zaccagnini, *Goldbach Variations: problems with prime numbers*, L'Educazione Matematica, Anno XXI, Serie VI **2** (2000), 47–57, http://people.math.unipr.it/alessandro.zaccagnini/psfiles/papers/Goldbach_E.pdf.
[54] A. Zaccagnini, *A note on the sum of a prime and a polynomial*, Quart. J. Math. Oxford **52** (2001), 519–524.

Prof. Alessandro Zaccagnini
Dipartimento di Matematica
Università degli Studi di Parma
Parco Area delle Scienze, 53/a
43124 Parma, ITALIA
Tel. 0521 906902 – Telefax 0521 906950
e-mail: alessandro.zaccagnini@unipr.it
web page: http://people.math.unipr.it/alessandro.zaccagnini/