Some recent developments at the intersection of Diophantine approximation, analytic number theory, and irregularities of distributions

Christoph Aistleitner, Roswitha Hofer and Gerhard Larcher

1 Introduction

The purpose of the present paper is to exhibit some recent developments at the intersection of metric Diophantine approximation and the theory of almost everywhere convergence of function series, analytic number theory, and the theory of irregularities of distributions modulo one. The connection between these diverse areas is made by so-called GCD sums (sums involving greatest common divisors), which will be discussed in detail below. These sums are of the form

$$\frac{1}{N} \sum_{k,l=1}^{N} \frac{(\gcd(n_k, n_l))^{2\alpha}}{(n_k n_l)^{\alpha}},$$

(1)

or, more generally,

$$\sum_{k,l=1}^{N} c_k \overline{c}_l \frac{(\gcd(n_k, n_l))^{2\alpha}}{(n_k n_l)^{\alpha}}.$$

(2)

Here \(n_1, \ldots, n_N\) are distinct positive integers, \(\alpha\) is a real parameter, which usually is from the range \([1/2, 1]\), and \(c_k\) are (real or complex) coefficients which are normalized such that \(\sum |c_k|^2 = 1\). The significance of these sums (in the case \(\alpha = 1\)) was probably first observed by Koksma in the 1930s. The problem of finding the maximal possible order (over all configurations \(n_1, \ldots, n_N\)) of the sum in (1) for \(\alpha = 1\) was suggested as a prize problem to the Wiskundig Genootschap at Amsterdam by Erdős, and was shortly later solved by Gál [16], who proved that this sum is of order at most \((\log \log N)^2\) and that this upper bound is optimal. The problem in the case \(\alpha = 1/2\) appears in the context of the Duffin–Schaeffer conjecture, a notoriously difficult open problem in Diophantine approximation, in a paper of Dyer and Harman [14].
However, the maximal order of the GCD sum in this case was only found very recently by Bondarenko and Seip [10, 11], where it was shown to be

$$\exp\left(c \frac{\sqrt{\log N} \sqrt{\log \log \log N}}{\sqrt{\log \log N}}\right).$$

The intermediate case $\alpha \in (1/2, 1)$ was solved in [2]; the maximal order in this case is

$$\exp\left(c_{\alpha} \frac{(\log N)^{1-\alpha}}{\left(\log \log N\right)^{\alpha}}\right).$$

(3)

For $\alpha > 1$ the optimal order is easily seen to be at most $c_{\alpha}$, and in the case $\alpha \in (0, 1/2)$ an optimal solution was found in [9]. Thus the problem concerning the maximal order of GCD sums is now completely solved. The upper bounds for (2) are of the same order as those for (1). The problem of finding the maximal value (also over the coefficients $c_1, \ldots, c_N$) of the sum in (2) has a natural interpretation in terms of finding the largest eigenvalue of a so-called GCD matrix, a problem which is of some interest in its own right. See [3, 21].

In the subsequent sections we will show how GCD sums arise in different areas of mathematics and which role they play there. However, before moving on we want to make a few remarks on where the upper bounds mentioned above come from, how they are obtained, and on a closely related similar problem.

Gál’s proof was based on a combinatorial optimization argument: if a configuration gives the maximal value for the GCD sum, then it must have strong structural properties (since otherwise the value of the GCD may be further increased). Such structural properties are for example the fact that for every number contained in \(\{n_1, \ldots, n_N\}\), all its divisors must be contained as well. Another such structural property is the fact that there cannot be too many primes involved when factorizing all the numbers \(\{n_1, \ldots, n_N\}\). As a hint as to where (3) comes from, assume that $N = 2^M$ and let \(\{n_1, \ldots, n_N\}\) be the set of all square-free numbers generated by the first $M$ primes $p_1, \ldots, p_M$. Then the GCD sum in (1) has a strong symmetric structure, and one can calculate

$$\frac{1}{N} \sum_{k,l=1}^{N} \frac{(\gcd(n_k, n_l))^{2\alpha}}{(n_k n_l)^{\alpha}} = \prod_{m=1}^{M} (1 + p_m^\alpha).$$

(4)

This product is easily shown to be of order roughly (3).
The upper bound given in [2] for the case $\alpha \in (1/2, 1)$ uses a strategy similar to Gál's at the beginning, but involves some heave machinery from complex analysis (analysis on the infinite-dimensional polydisc, to be more precise). The case $\alpha = 1/2$ is even more difficult. Quite amazingly, an alternative, totally different proof was given for the case $\alpha \in (1/2, 1)$ in a paper of Lewko and Radziwiłł [20]; they used an interpretation of the GCD sum as an integral involving a random model for the Riemann zeta function. This proof indicates a connection between GCD sums and properties of the Riemann zeta function, a point which will be discussed in more detail in Section 3 below.

Finally we mention the problem of maximizing the expression

$$\sum_{k,l=1}^{N} c_k \overline{c_l} \frac{(\gcd(k, l))^{2\alpha}}{(kl)^{\alpha}},$$

where $c_k$ are normalized such that $\sum |c_k|^2 = 1$ (and where the maximization problem is over all coefficients $c_1, \ldots, c_N$). Note that the maximal value of this expression is obviously dominated by the one in (2). The problem was solved by Hilberdink [17], who obtained the upper bounds

$$c (\log \log N)^2 \quad \text{for } \alpha = 1,$$

$$\exp \left( c \alpha \frac{\log N}{\log \log N} \right) \quad \text{for } \alpha \in (1/2, 1),$$

$$\exp \left( c \frac{\sqrt{\log N}}{\sqrt{\log \log N}} \right) \quad \text{for } \alpha = 1/2.$$

Note that these upper bounds are similar to those given above, but different for $\alpha \in (1/2, 1)$ and $\alpha = 1/2$. However, there is a close similarity in the problems of finding numbers $n_1, \ldots, n_N$ maximizing (1) and finding coefficients $c_1, \ldots, c_N$ maximizing (5), respectively. For instance, to find an example achieving (6) one can take coefficients supported on the square-free integers generated by the first $M$ primes and make a calculation similar to (4), where however one has to choose $M \approx \log N / \log \log N$ rather than $M \approx \log N$ to make sure that the largest non-zero coefficient really has index at most $N$.

2 Metric Diophantine approximation and almost everywhere convergence of function series

To see how GCD sums appear in metric number theory, assume that $f$ is the indicator function of an interval modulo one, centered such that $\int_0^1 f(x) \, dx =$
0, and that we want to calculate

$$
\int_0^1 \left( \sum_{k=1}^{N} c_k f(n_k x) \right)^2 \, dx,
$$

(7)

where $c_k$ are real coefficients and $(n_k)_{k \geq 1}$ is a sequence of distinct positive integers. This is a very natural problem, since this integral is just the variance of the sum $\sum_{k=1}^{N} c_k f(n_k x)$, understood as a random variable on $([0,1], \mathcal{B}(0,1), \lambda)$, where $\lambda$ is the Lebesgue measure. An upper bound for this variance, together with Markov’s inequality and the Borel-Cantelli lemma, will allow us to make metric statements on the asymptotic order (or convergence/divergence) of the sum as $N \to \infty$.

Let $f(x) \sim \sum_{j=-\infty}^{\infty} a_j e^{2\pi i j x}$ be the Fourier series of $f$. By orthogonality, the integral in (7) equals

$$
\sum_{1 \leq k, l \leq N, j_1, j_2 \in \mathbb{Z}, j_1 n_k = j_2 n_l} c_k c_l a_{j_1} a_{j_2}.
$$

We assumed that $f$ is the indicator function of an interval, which allows us to deduce that $|a_j| \leq (\pi |j|)^{-1}$. Since $f(x)$ has mean zero, we have $a_0 = 0$. Thus (7) is dominated by

$$
\frac{1}{\pi^2} \sum_{1 \leq k, l \leq N, j_1, j_2 \in \mathbb{Z}\setminus\{0\}, j_1 n_k = j_2 n_l} \frac{c_k c_l}{j_1 j_2}.
$$

Now the crucial observation is that $j_1 n_k = j_2 n_l$ whenever $j_1 = j n_l / \gcd(n_k, n_l)$ and $j_2 = j n_k / \gcd(n_k, n_l)$ for some integer $j$. Thus we get the upper bound

$$
\frac{1}{\pi^2} \sum_{1 \leq k, l \leq N} \sum_{j \in \mathbb{Z}\setminus\{0\}} \frac{c_k c_l (\gcd(n_k, n_l))^2}{j^2 n_k n_l} = \frac{1}{3} \sum_{1 \leq k, l \leq N} \frac{c_k c_l (\gcd(n_k, n_l))^2}{n_k n_l},
$$

which gives an expression such as the one in (2). The same reasoning works if $f$ is not an indicator function of an interval, but a function of bounded variation [18]. Additionally, if one knows that the Fourier coefficients $a_j$ with indices $j$ close to zero are small (for example because $f$ is the indicator of a short interval) and wants to exploit this fact, then one is led quite naturally to a GCD sum with parameter $\alpha = 1/2$. GCD sums with $\alpha \in (1/2, 1)$ are obtained for example by interpolation between the cases $\alpha = 1$ and $\alpha = 1/2$ using a weighted geometric mean. As noted, arguments of this form play
a major role in a paper of Dyer and Harman [14] on the Duffin–Schaeffer conjecture. In [2] and [20] such arguments allowed the authors to solve a decades-old open problem on the almost everywhere (a.e.) convergence of series of dilated functions, which can be seen as a generalization of the problem asking for the a.e. convergence of Fourier series, which was famously solved by Carleson [12] in 1966. Currently, the first and third author of the present paper are preparing a manuscript together with Mark Lewko on the metric theory of pair correlations, where GCD sums will also play a crucial role and will lead to improvements of results such as those obtained by Rudnick and Zaharescu [23].

3 Analytic number theory

Quite unexpectedly, recently a connection between GCD sums and certain properties of the Riemann zeta function was established. One indication of such a connection was exposed by the proof of Lewko and Radziwill of upper bounds for the maximal value of GCD sums, which, as noted above, is based on an interpretation of GCD sums in terms of a random model for the Riemann zeta function. However, a direct formal connection was established in a paper of Hilberdink [17], who modified the resonance method of Soundararajan [24] in such a way that GCD sums show up there in a natural way. In the sequel, we want to describe this connection along general lines.

A well-known conjecture concerning the Riemann zeta function is the Lindelöf hypothesis, which asserts that \( \zeta(1/2 + it) = O(t^\varepsilon) \) for all \( \varepsilon > 0 \). The Lindelöf hypothesis is far from being proved; the exponent 1/6 (rather than \( \varepsilon \)) in the upper bound, due to Hardy–Littlewood, has only been slightly improved during a century. The Lindelöf hypothesis is weaker than the Riemann hypothesis, whose truth would imply that

\[
\zeta(\sigma + it) = O \left( \exp \left( \frac{c_{\sigma}(\log t)^{2-2\sigma}}{\log \log t} \right) \right)
\]

for fixed \( \sigma \in [1/2, 1) \). The best known lower bounds are

\[
\zeta(\sigma + it) = \Omega \left( \exp \left( \frac{c_{\sigma}(\log t)^{1-\sigma}}{(\log \log t)^{\sigma}} \right) \right)
\]

for \( \sigma \in (1/2, 1) \), due to Montgomery [22] (and conjectured to be optimal), and

\[
\zeta(1/2 + it) = \Omega \left( \exp \left( \frac{c_{\sigma} \sqrt{\log t} \sqrt{\log \log \log t}}{\sqrt{\log \log t}} \right) \right),
\]
which was recently shown by Bondarenko and Seip [10] using some of the methods described in this section.

Suppose we want to establish a lower bound for the Riemann zeta function. The idea of the resonance method is to find a function $A(t)$ such that

$$ I_1 := \int_0^T |\zeta(\sigma + it)A(t)|^2 \, dt $$

is “large” and

$$ I_2 := \int_0^T |A(t)|^2 \, dt $$

is “small”, since obviously there must exist a value of $t \in [0, T]$ for which $|\zeta(\sigma + it)|^2$ is at least as large as the quotient $I_1/I_2$. In Soundararajan’s original argument the integral $I_1$ shows up with exponent 1 rather than 2; the version with exponent 2 and the following observations are due to Hilberdink [17]. Assume that we can approximate $\zeta$ by a Dirichlet polynomial

$$ \zeta(\sigma + it) \approx \sum_{n \leq N} \frac{1}{n^{\sigma + it}}, $$

which is just the initial segment of its representation as a Dirichlet series. Assume that $A(t)$ also is a Dirichlet polynomial of the form

$$ A(t) = \sum_{k=1}^K b_k k^{it}. $$

Then, just by squaring out, we have

$$ I_1 \approx \int_0^T \sum_{1 \leq m,n \leq N} \sum_{1 \leq k,l \leq K} \frac{b_k \overline{b_l}}{(mn)^\sigma} \left( \frac{mk}{nl} \right)^{it} \, dt $$

$$ = \int_0^T \sum_{1 \leq m,n \leq N} \sum_{1 \leq k,l \leq K} \frac{b_k \overline{b_l}}{(mn)^\sigma} \left( \frac{mk}{nl} \right)^{it} \, dt \quad (8) $$

$$ + \int_0^T \sum_{1 \leq m,n \leq N} \sum_{1 \leq k,l \leq K} \frac{b_k \overline{b_l}}{(mn)^\sigma} \left( \frac{mk}{nl} \right)^{it} \, dt. \quad (9) $$

In line (8) we have $(mk/nl)^{it} = 1$ for all $t$, and thus this line equals

$$ T \sum_{1 \leq m,n \leq N} \sum_{1 \leq k,l \leq K} \frac{b_k \overline{b_l}}{(mn)^\sigma}. $$
We have \( mk = nl \) whenever \( m = jl/(\gcd(k, l)) \) and \( m = jk/(\gcd(k, l)) \) for some integer \( j \), so this sum is of order

\[
T \sum_{k,l=1}^{K} b_k \overline{b}_l \frac{(\gcd(k,l))^{2\sigma}}{(kl)^{\sigma}}
\]

(if we ignore the values of \( j \) above), and thus we have in a very natural way obtained a GCD sum as in (5). On the other hand, it turns out that the term in line (9) is small if \( N \) and \( K \) are small powers of \( T \), and that \( I_2 \) is of order \( T \) also if \( K \) is a small power of \( T \). Thus a lower bound for the GCD sum yields a lower bound for the maximum of the Riemann zeta function. Furthermore, the argument around line (4) suggests how a “good” resonator \( A(t) \) could be constructed, namely in a multiplicative form as a finite Euler product. In comparison it is remarkable how in the argument in the present section the functions \( (m/n)^{it} \), which are “almost orthogonal” on \([0, T]\) if \( m, n \) are not too large, play the role of the (orthogonal) trigonometric system in the previous section.

The purpose of the present section was only to exhibit the way how GCD sums arise in the context of the Riemann zeta function, and can be used to prove the existence of large values of the zeta function. The argument in the lines above corresponds to the sum in (5) and gives the lower bounds for such GCD sums mentioned in the lines below equation (5). In comparison to what was said at the beginning of the present section, these lower bounds are weaker than the best ones known for large values of the zeta function. This defect can be overcome by using an “extremely long resonator”, which leads to a GCD sum as in (2) rather than (5). The problem with this long resonator is that one loses the “almost orthogonality” property which played a crucial role in the argument sketched above. To see how this problem can be solved, we refer the reader to [1, 10]. There are further issues when trying to generalize this method to \( L \)-functions; see [7].

4 Irregularities of distributions

A sequence of real numbers \((x_n)_{n \geq 1}\) in \([0, 1]\) is called uniformly distributed modulo one (u.d. mod 1) if

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} 1_{[a,b]}(x_n) = b - a
\]
for all \([a, b] \subset [0, 1]\). Here, and in the sequel, \(1_{[a,b]}\) denotes the indicator function of the interval \([a, b]\). The most classical example in this theory is the sequence of fractional parts \((\{n\alpha\})_{n \geq 1}\), which is u.d. mod 1 if and only if \(\alpha \notin \mathbb{Q}\), which was shown independently by Bohl, Sierpiński and Weyl in 1909/1910. Another famous example (Weyl [25], 1916) states that for distinct integers \((n_k)_{k \geq 1}\) the sequence \((\{n_k\alpha\})_{k \geq 1}\) is u.d. mod 1 for almost all \(\alpha\) in the sense of Lebesgue measure — however, in this general setting it is usually totally impossible to explicitly determine the exceptional set of measure zero.

Uniform distribution modulo one can be quantified using the notion of the discrepancy. Let \(x_1, \ldots, x_N \in [0, 1]\). Then the discrepancy of these numbers is defined as

\[
D_N(x_1, \ldots, x_N) = \sup_{[a,b] \subset [0,1]} \left| \frac{1}{N} \sum_{n=1}^{N} 1_{[a,b]} - (b - a) \right|.
\]

An infinite sequence is u.d. mod 1 if and only if the discrepancy of its first \(N\) elements tends to zero as \(N \to \infty\). (Note the similarity of this notion to the Glivenko–Cantelli theorem in probability theory). A quantitative version of Weyl's theorem was given by R.C. Baker [8]: for every strictly increasing sequence of integers \((n_k)_{k \geq 1}\) we have

\[
D_N(\{n_1\alpha\}, \ldots, \{n_N\alpha\}) = \mathcal{O}\left(\frac{(\log N)^{3/2+\varepsilon}}{\sqrt{N}}\right) \quad \text{a.e.} \quad (10)
\]

This result is known to be optimal, except for the exponent of the logarithmic term (whose optimal value is a major open problem in metric number theory). The proof makes heavy use of Carleson's theorem (in the form of the Carleson–Hunt inequality), as well as of the Erdős–Turán inequality, which allows one to estimate the discrepancy in terms of exponential sums. The method of proof is similar to the one mentioned in Section 2: one first establishes upper bounds for the variance and then uses Markov's inequality and the Borel–Cantelli lemma.

Proving metric lower bounds is a totally different business; simply speaking, the problem with lower bounds is that they can neither be deduced from moment bounds nor with the (second) Borel–Cantelli lemma, because large moments do not necessarily imply large exceptional sets and because the second Borel–Cantelli lemma is not applicable since there is no independence. Recently the authors of the present paper have developed a new, general method which allows to prove lower bounds in metric discrepancy theory,
which makes crucial use of GCD sums. Let \((n_k)_{k \geq 1}\) be a sequence of integers. By the so-called Koksma inequality (see for example [13, 19], which are the standard references for uniform distribution theory and discrepancy theory) we have

\[
ND_N \left( \{n_1 \alpha\}, \ldots, \{n_N \alpha\} \right) \geq \frac{1}{4h} \left| \sum_{k=1}^{N} e^{2\pi i h n_k \alpha} \right|,
\]

where \(h\) is an arbitrary positive integer. This implies that

\[
ND_N \left( \{n_1 \alpha\}, \ldots, \{n_N \alpha\} \right) \geq \frac{1}{4N^{2 \varepsilon}} \sum_{1 \leq h \leq N^{\varepsilon}} \left| \sum_{k=1}^{N} e^{2\pi i h n_k \alpha} \right|,
\]

(11)

where \(\varepsilon\) is a small real number. Let \(A\) denote a set of those \(\alpha \in [0, 1]\) where \(\left| \sum_{k=1}^{N} e^{2\pi i n_k \alpha} \right|\) is of size at least \(N^{c+\eta}\), where \(c\) is an appropriate real constant (coming from the \(L^1\) norm of the exponential sum) and \(\eta\) is an appropriate real parameter. Then the right-hand side of (11) is certainly large whenever

\[
\sum_{1 \leq h \leq N^{\varepsilon}} 1_A(\{h \alpha\}) \approx N^{\varepsilon} \lambda(A),
\]

(12)

is large. Thus we have a problem in Diophantine approximation, and have to check how regularly the fractional parts \(\{h \alpha\}\) are contained or not contained in \(A\) as \(h\) takes the values \(1, 2, \ldots, N^{\varepsilon}\). For this we have to estimate the variance of (12), which can be done as in Section 2 and for which some information on the regularity of the set \(A\) is required. Then we have

\[
\sum_{1 \leq h \leq N^{\varepsilon}} 1_A(\{h \alpha\}) \approx N^{\varepsilon} \lambda(A),
\]

for “typical” \(\alpha\), where \(\lambda\) denotes the Lebesgue measure, provided that the variance is not too large (of significantly smaller order than \(N^{\varepsilon}\), which turns out to be the case). If we know for the \(L^1\) norm of the exponential sum that \(\left| \sum_{k=1}^{N} e^{2\pi i n_k \alpha} \right| \approx N^c\), then (by dyadic splitting and the pigeon hole principle) the measure of \(A\) has to be roughly \(N^{-\eta}\) for some appropriate \(\eta\), and we have

\[
ND_N \left( \{n_1 \alpha\}, \ldots, \{n_N \alpha\} \right) \gg N^{-2 \varepsilon} N^{c+\eta} N^{\varepsilon} N^{-\eta} = N^{c-\varepsilon}.
\]

Refining the argument a bit and using the (first) Borel–Cantelli lemma gives an asymptotic result for \(N \to \infty\). Thus lower bounds for \(L^1\) norms of exponential sums together with upper bounds on GCD sums lead to lower bounds in metric discrepancy theory. This is a novel method, which has led to several interesting results. Among them are the following.
Results for \((n_k)_{k \geq 1}\) being the so-called Thue–Morse sequence of integers, that is the sequence of positive integers having even sum-of-digits in base 2. The necessary estimates for \(L^1\) norms of exponential sums come from a paper of Fouvry and Mauduit [15]. The new results in [4] show an interesting deviation between the metric behavior of exponential sums and that of the discrepancy, respectively, something that has not been observed before.

Results for \((n_k)_{k \geq 1}\) being the values \(p(k)\) of a polynomial \(p \in \mathbb{Z}[x]\). The necessary \(L^1\) bounds come from bounds for the number of representations of an integer as the difference of two values of the polynomial. See [6].

By a classical trick, which is based on a clever application of Hölder's inequality, a lower bound for the \(L^1\) norm of exponential sums follows from an upper bound on the \(L^4\) norm. Moreover, the \(L^4\) norm of an exponential sum of \((n_k \alpha)_{1 \leq k \leq N}\) is a purely combinatorial object, and actually is equal to what is called the additive energy in additive combinatorics (a notion which has received a lot of attention recently). Thus upper bounds for the additive energy imply lower bounds for the metric discrepancy. In particular, minimal additive energy (of order \(N^2\)) implies an \(L^1\) norm of maximal order (that is, \(\sqrt{N}\)), which, by the argument above (for \(c = 1/2\)) gives \(ND_N \approx \sqrt{N}\). In view of Baker's result (10), this is essentially the maximal order of the metric discrepancy. In other words, results from additive combinatorics allowed us to identify several new classes of integer sequences \((n_k)_{k \geq 1}\) which give the maximal possible order in metric discrepancy theory. See [5] for details.

Acknowledgements

The first author is supported by the Austrian Science Fund (FWF), projects I1751-N26 and Y-901-N35. The first and third author are supported by the FWF project F5507-N26 and the second author is supported by the FWF project F5505-N26, which are both parts of the Special Research Program Quasi-Monte Carlo Methods: Theory and Applications.

References


Christoph Aistleitner, Roswitha Hofer, Gerhard Larcher
Department of Financial Mathematics and Applied Number Theory
Altenbergerstraße 69
4040 Linz, Austria

*E-mail addresses:* christoph.aistleitner@jku.at, roswitha.hofer@jku.at, gerhard.larcher@jku.at