Exponential Sums in Coding Theory and Cryptology

Igor E. Shparlinski
Macquarie University

igor@comp.mq.edu.au

Lecture Notes of Tutorial Lectures given at the
Institute for Mathematical Sciences of the NUS

Singapore, 23–26 July, 2001

The author would greatly appreciate any comments,
corrections and critical remarks

Contents

Introduction

In these lecture notes we will try to exhibit, in a very informal way, some useful and sometimes surprising relations between exponential sums, which is a celebrated tool on analytical number theory, and several important problems of such applied areas as coding theory and cryptology.
One can certainly ask two natural questions:

- Why Exponential Sums?

  And the answer is:

  - Because they are beautiful and I like them;
  - Because exponential sums allow us to show the existence of objects with some special properties.

- Why Coding Theory and Cryptology?

  - Because they are beautiful and I like them as well;
  - Because to design/analyze some codes and cryptographic schemes we need to find objects with some special properties:
    - "*good*" for designs;
    - "*bad*" for attacks.

The main goal of these lectures is to show that exponential sums are very useful, yet user friendly objects, provided you know how to approach them.
I will also provide a necessary background for everybody who would like to learn about this powerful tool and to be able to use it in her and his own work. I do not pretend to give a systematic introduction to the subject but rather I intend help to get started in making exponential sums an active working tool, at least in the situation where their application does not require any sophisticated technique or advanced analytical methods. I hope my lectures should help to develop some feeling of the kinds of questions where exponential sums can be useful and if you see that the actual application is beyond your level of expertise you can always seek an advise from one of the numerous experts in number theory (who probably otherwise would never know about your problem).
It is well know that for many years number theory was the main area of applications of exponential sums. Such applications include (but are not limited to)

- Uniform distribution (H. Weyl);
- Additive problems such the Goldbach and Waring problems (G. H. Hardy, J. E. Littlewood, R. Vaughan, I. M. Vinogradov);
- Riemann zeta function and distribution of prime numbers (J. Littlewood, N. M. Korobov, Yu. V. Linnik, E. C. Titchmarsh, I. M. Vinogradov).

However it has turned out that exponential sums provide a valuable tool for a variety of problems of theoretical computer science, coding theory and cryptography, see [65, 66].
I will try to explain:

- What we call exponential sums.
- How we estimate exponential sums (and why we need this at all).
- What is current state of affairs.
- What kind of questions can be answered with exponential sums.
- How various cryptographic and coding theory problems lead to questions about exponential sums.

Unfortunately there is no systematic textbook on exponential sums. However one can find a variety of results and applications of exponential sums in [33, 45, 65, 75].

Acknowledgment. The author would like to thank Harald Niederreiter for careful reading of the manuscript and numerous helpful suggestions.

Chapter 1

Exponential Sums — Basic Notions

1.1 Getting Started

1.1.1 Exponential Sums — What Are They?

Exponential sums are objects of the form

$$S(\mathcal{X}, F) = \sum_{x \in \mathcal{X}} \mathbf{e}(F(x))$$

where

$$\mathbf{e}(z) = \exp(2\pi i z),$$

$\mathcal{X}$ is an arbitrary set, F is a real-valued function on $\mathcal{X}$.
In fact $\mathcal{X}$ could be a set of vectors, in this case we talk about **multiple sums**.

1.1.2 Exponential Sums — What Do We Want From Them?

Certainly it would be very good to have a closed form expression for the sums $S(\mathcal{X}, F)$. Unfortunately there very few examples when we have such formulas. On the other hand, for main applications of exponential sums we do not need to know $S(\mathcal{X}, F)$ exactly. It is quite enough to have an **upper bound** on $S(\mathcal{X}, F)$, which is the main task of this area.
First of all we remark that because $|\mathbf{e}(z)| = 1$ for every real z,

$$|S(\mathcal{X}, F)| \leq \#\mathcal{X}.$$

This is the **trivial bound**.

We are interested in getting stronger bounds. Of course, to be able to prove such a bound we need some conditions on $\mathcal{X}$ and F. For example, if F is an integer-valued function then $\mathbf{e}(F(x)) = 1$ and $S(\mathcal{X}, F) = \#\mathcal{X}$.

1.1.3 Exponential Sums — How Do We Classify Them?

There are exponentially many different types of exponential sums.

If $\mathcal{X}$ is a set of vectors, we talk about **multiple sums**. In particular in the two-dimensional case we talk about **double sums**. Double sum technique provides an invaluable tool in estimating one-dimensional sums.

A very important class of exponential sums consists of **rational sums**. Those are the sums with functions F of the form $F(x) = f(x)/m$ where $f : \mathcal{X} \to \mathbb{Z}$ is an integer-valued function on $\mathcal{X}$. The number m is called the **denominator** of the exponential sum $S(\mathcal{X}, F)$.

It is convenient to introduce one more notation

$$\mathbf{e}_m(z) = \exp(2\pi i z/m)$$

(thus $\mathbf{e}_1(z) = \mathbf{e}(z)$). Therefore we have

$$S(\mathcal{X}, F) = \sum_{x \in \mathcal{X}} \mathbf{e}_m(f(x)).$$

1.2 Timeline

Exponential sums are almost 200 years old. It is a long history of triumphs and disappointments. Below I tried to outline some most important events of this dramatic history.

1.2.1 Johann Carl Friedrich Gauss, 1811

Exponential sums were introduced to number theory by **Gauss** in [19]. The sums he introduced and studied

$$G(a, m) = \sum_{x=0}^{m-1} e_m(ax^2)$$

are called "Gaussian sums" in his honor. Sometimes this name is extended to more general sums

$$G_n(a, m) = \sum_{x=0}^{m-1} e_m(ax^n)$$

as well. Gaussian sums $G(a, m)$ is one of very few examples when one can actually evaluate exponential sums explicitly. It should be noticed that the way Gauss used these sums is very different from modern applications of exponential sums.

1.2.2 Hermann Klaus Hugo Weyl, 1916

Hermann Weyl was probably the first mathematician who understood the great power and potential of this method. Besides creating the first general method of bounding exponential sums [80], he also found very important connections with uniform distribution of sequences which underlie many further applications of this method.

1.2.3 Godfrey Harold Hardy and John Edensor Littlewood, 1920

Godfrey Hardy and John Littlewood [24] found new applications of exponential sums to some very important number theoretic problems and invented their "circle method" which is now routinely used for a large number of applications [75]. John Littlewood [46] also introduced exponential sums in studying the Riemann zeta function.

1.2.4 Louis Joel Mordell, 1932

Louis Mordell [51] created a new method of estimating rational exponential sums with polynomials with prime denominator. Despite that the method is obsolete and superseded by the Andre Weil method [79], it exhibited some very important principles and is has not lost its value as a teaching tool in the theory of exponential sums.

1.2.5 Ivan Matveevich Vinogradov, 1935

Ivan Vinogradov developed a principally new method of estimating general exponential sums with polynomials with irrational coefficients [77] (much stronger that H. Weyl's method) and also the method of bounding exponential sums where the set $\mathcal{X}$ consists of prime numbers of a certain interval [78]. He obtained extremely strong results for such classical problem as the *Waring problem* and the *Goldbach problem* and the bounds for the zeros of the Riemann zeta function. Even now, 65 years later we do not have anything essentially stronger.

1.2.6 Loo-Keng Hua, 1947

Loo-Keng Hua [32] created a new method of estimating rational exponential sums with arbitrary denominator. The method is based on Chinese Remainder Theorem to reduce the general case to the case of prime power denominator, and then using a kind of Hensel lifting to reduce the case of prime power denominator to the case of prime denominator. Almost all works on exponential sums with arbitrary denominator follow this pattern.

1.2.7 Andre Weil, 1948

Andre Weil [79] invented an algebraic-geometry method of estimating "rational" exponential sums with prime denominator. In many case the result are close to best possible. It still remains the most powerful tool in this area.

1.2.8 Pierre Deligne, 1972

Pierre Deligne [15] has obtained a very important extension of the algebraic geometry method to bounds of multiple sums with polynomials and rational functions with prime denominator.

1.2.9 You, ????

There also have been many other exceptional researchers and outstanding results and methods but no " breakthroughs". An excellent outline of older results is given by Loo-Keng Hua [33]. Maybe its **your** turn now! The area deserves your attention.

1.3 Some Terminology

1.3.1 Rational Exponential Sums

We concentrate on the simplest, yet most useful, well-studied and attractive class of **rational exponential sums**. That is, the function $F(x) = f(x)/m$ takes rational values with integer denominator $m > 1$.

In fact very often we concentrate only on the case of prime denominators. Sometimes it is convenient to think that $f(x)$ is defined on elements of the finite field $\mathbb{F}_p$ of p elements.

Examples:

- $F(x) = f(x)/p$ where f is a polynomial with integer coefficients (alternatively one can think that f is a polynomial with coefficients from $\mathbb{F}_p$);
- $F(x) = g^x/p$ where $g > 1$ is an integer (alternatively one can think that $g \in \mathbb{F}_p$).

1.3.2 Complete and Incomplete Exponential Sums

Very often the function $f(x)$ in $F(x) = f(x)/m$ is purely periodic modulo m with period T. Then the sum

$$S(f) = \sum_{x=1}^{T} \mathbf{e}_m(f(x))$$

is called a **complete sum**.
A shorter sums

$$S(f,N) = \sum_{x=1}^{N} \mathbf{e}_m(f(x))$$

with $1 \le N \le T$ is called an **incomplete sum**.
Examples:

- If $f(x)$ a polynomial with integer coefficients then it is periodic modulo p with period p;
- $f(x) = g^x$ where $g > 1$ is an integer with $\gcd(g,p) = 1$ then it is periodic modulo p with period t where t is the multiplicative order of g modulo p.

Typically, incomplete sums (especially when N is relatively small to T) are much harder to estimate.

Chapter 2

Simplest Bounds and Applications

2.1 The Basic Case — Linear Sums

Certainly the simplest (and easiest) exponential sums one can think of are **linear exponential sums**, that is, exponential sums with

$$F(x) = ax/p.$$

The following simple results give a complete description of such sums (a very unusual situation ...). It provides a very good warming up exercise.

Theorem 1

$$\sum_{x=0}^{m-1} \mathbf{e}_m(ax) = \begin{cases} 0, & \text{if } a \not\equiv 0 \pmod{m}, \\ m, & \text{if } a \equiv 0 \pmod{m}. \end{cases}$$

Proof. The case $a \equiv 0 \pmod{m}$ is obvious because each term is equal to 1. The case $a \not\equiv 0 \pmod{m}$ ... is obvious as well, because it is a sum of a geometric progressions with denominator $q = \mathbf{e}_m(a) \neq 1$ thus

$$\sum_{x=0}^{m-1} \mathbf{e}_m(ax) = \sum_{x=0}^{m-1} q^x = \frac{q^m - 1}{q - 1} = \frac{\mathbf{e}_m(ma) - 1}{\mathbf{e}_m(a) - 1} = \frac{1 - 1}{\mathbf{e}_p(a) - 1} = 0.$$

□

2.2 Nice Result Almost for Free

The following statement is a very instructive example showing the great power of the exponential sum method. The result is a rather **nontrivial** statement which follows immediately from **trivial** Theorem 1. In fact I am not aware of any alternative proof of this statement whose formulation has nothing to do with exponential sums.

Let $\mathcal{X}$ be **any** set of $\mathbb{Z}$ and let f be function $f : \mathcal{X} \to \mathbb{F}_p$.

Let $N_k(a)$ be the number of solutions of

$$f(x_1) + \ldots + f(x_k) \equiv f(x_{k+1}) + \ldots + f(x_{2k}) + a \pmod p.$$

where $x_1, \ldots, x_{2k} \in \mathcal{X}$ and a is an integer.

Theorem 2 $N_k(a) \le N_k(0)$.

Proof. By Theorem 1

$$\begin{aligned} N_k(a) = \sum_{x_1,\ldots,x_{2k}\in\mathcal{X}} \frac{1}{p}\sum_{c=0}^{p-1} \mathbf{e}_p\Big(c\big(f(x_1) + \ldots + f(x_k) \\ -f(x_{k+1}) - \ldots - f(x_{2k}) - a\big)\Big). \end{aligned}$$

Rearranging,

$$N_k(a) = \frac{1}{p}\sum_{c=0}^{p-1} \mathbf{e}_p(-ca)\left(\sum_{x\in\mathcal{X}} \mathbf{e}_p\left(cf(x)\right)\right)^k \left(\sum_{x\in\mathcal{X}} \mathbf{e}_p\left(-cf(x)\right)\right)^k.$$

Because for any real u,

$$\mathbf{e}_p(-u) = \overline{\mathbf{e}_p(u)}$$

and for any complex z,

$$z\overline{z} = |z|^2,$$

we obtain

$$\begin{aligned} N_k(a) &= \frac{1}{p}\sum_{c=0}^{p-1} \mathbf{e}_p(-ca)\left|\sum_{x\in\mathcal{X}} \mathbf{e}_p\left(cf(x)\right)\right|^{2k} \\ &\le \frac{1}{p}\sum_{c=0}^{p-1} \left|\sum_{x\in\mathcal{X}} \mathbf{e}_p\left(cf(x)\right)\right|^{2k} = N_k(0). \end{aligned}$$

□

It is obvious that

$$\sum_{a=0}^{p-1} N_k(a) = \#\mathcal{X}^{2k}.$$

Indeed, any $2k$-tuple $(x_1, \dots, x_{2k}) \in \mathcal{X}^{2k}$ corresponds to one and only one congruence and will be counted exactly once.

Exercise 3 *Using Theorem 2 and the previous observation, prove that*

$$N_k(0) \geq \frac{\#\mathcal{X}^{2k}}{p}.$$

As you can see, Theorem 2 follows from the explicit expression of $N_k(a)$ via exponential sums. It also gives a lower bound on $N_k(0)$. Now we show that having some extra information about exponential sums involved in this expression one can show that all values of $N_k(a)$ are close to their expected value $\#\mathcal{X}^{2k}/p$.

In the formula

$$N_k(a) = \frac{1}{p} \sum_{c=0}^{p-1} \mathbf{e}_p(-ca) \left| \sum_{x \in \mathcal{X}} \mathbf{e}_p\left(cf(x)\right) \right|^{2k}$$

the term corresponding to $c = 0$ is $\#\mathcal{X}^{2k}/p$. Assume that we know a **non-trivial** upper bound

$$\max_{1 \leq c \leq p-1} \left| \sum_{x \in \mathcal{X}} \mathbf{e}_p\left(cf(x)\right) \right| \leq \#X\Delta$$

with some $0 \leq \Delta < 1$. Then each of the other $p - 1$ terms is at most $\#\mathcal{X}^{2k}\Delta^{2k}$. Therefore

$$\left| N_k(a) - \frac{\#\mathcal{X}^{2k}}{p} \right| \leq \#\mathcal{X}^{2k}\Delta^{2k}$$

For some k we get $\Delta^{2k} < p^{-1}$ and we have an asymptotic formula.

The smaller the value of Δ, the smaller the value of k is needed. If $\Delta = p^{-\delta}$ one can take $k = \lfloor 1/2\delta \rfloor + 1$.

Moral:

1. The **expected value** of $N_k(a)$ is given by the term corresponding to $c = 0$.

2. The **error term** depends on the quality of our bound of exponential sums.

2.3 Linear Sums Once Again

In Theorem 1 the argument x runs through the whole field $\mathbb{F}_p$ of p elements. A natural question to ask is: What if we take shorter sums

$$T_a(h) = \sum_{x=0}^{h-1} \mathbf{e}_m(ax)$$

with $0 \le h \le p-1$?
It is still the sum of a geometric progression with denominator $q = \mathbf{e}_m(a) \ne 1$ thus

$$|T_a(h)| = \left| \frac{q^h - 1}{q-1} \right| \le \frac{2}{|q-1|}.$$

We have

$$\begin{aligned} |q-1| &= |\mathbf{e}_m(a) - 1| = |\exp(\pi i a/m) - \exp(-\pi i a/m)| \\ &= 2|\sin(\pi a/m)|. \end{aligned}$$

Let $1 \le a \le m-1$. Put $b = \min\{a, m-a\}$. Then

$$|\sin(\pi a/p)| = |\sin(\pi b/m)| \ge \frac{2b}{m}$$

because $\sin(\alpha) \ge 2\alpha/\pi$ for $0 \le \alpha \le \pi/2$.
Therefore

$$|T_a(h)| \le \frac{m}{2\min\{a, m-a\}}$$

for $1 \le a \le m-1$.
This immediately implies:

Theorem 4

$$\sum_{a=1}^{m-1} \left| \sum_{x=k}^{k+h-1} \mathbf{e}_m(ax) \right| = O(m \log m).$$

Proof. We have

$$\left| \sum_{x=k}^{k+h-1} \mathbf{e}_m(ax) \right| = \left| \mathbf{e}_m(ak) \sum_{x=0}^{h-1} \mathbf{e}_m(ax) \right| \le \frac{m}{2\min\{a, m-a\}}.$$

Therefore

$$\sum_{a=1}^{m-1} \left| \sum_{x=k}^{k+h-1} \mathbf{e}_m(ax) \right| = m \sum_{a=1}^{m-1} \frac{1}{2\min\{a, m-a\}} \le 2m \sum_{1 \le a \le m/2} \frac{1}{2a}$$

and the result follows. $\square$

2.4 Distribution of Functions Modulo p

Here we obtain the first general results illustrating how exponential sums can be used to gain some information about the distribution of functions modulo p.
Another interpretation of this result is a statement about the uniformity of distribution of the fractional parts

$$\left\{\frac{f(x)}{p}\right\}, \quad x \in \mathcal{X},$$

in the unit interval $[0,1]$.
Let k and $h \le p$ be integer. Denote

$$N_f(k,h) = \#\left\{x \in \mathcal{X} \,:\, f(x) \equiv v \pmod p,\ v \in [k, k+h-1]\right\}.$$

Theorem 5 *If*

$$\max_{1 \le c < p} \left| \sum_{x \in \mathcal{X}} \mathbf{e}_p\left(cf(x)\right) \right| \le \# X\Delta$$

then

$$\max_k \max_{0 \le h \le p-1} \left| N_f(k,h) - \frac{\#\mathcal{X}h}{p} \right| = O(\#X\Delta \log p).$$

Proof. We have

$$\begin{aligned} N_f(k,h) &= \sum_{x \in \mathcal{X}} \sum_{v=k}^{k+h-1} \frac{1}{p} \sum_{c=0}^{p-1} \mathbf{e}_p\left(c(f(x)-v)\right) \\ &= \frac{1}{p} \sum_{c=0}^{p-1} \sum_{x \in \mathcal{X}} \sum_{v=k}^{k+h-1} \mathbf{e}_p\left(-cv\right) \mathbf{e}_p\left(cf(x)\right) \\ &= \frac{\#\mathcal{X}h}{p} + \frac{1}{p} \sum_{c=1}^{p-1} \sum_{v=k}^{k+h-1} \mathbf{e}_p\left(-cv\right) \sum_{x \in \mathcal{X}} \mathbf{e}_p\left(cf(x)\right). \end{aligned}$$

Therefore

$$\begin{aligned} &\left| N_f(k,h) - \frac{\#\mathcal{X}h}{p} \right| \\ &\quad \le \frac{1}{p} \sum_{c=1}^{p-1} \left| \sum_{v=k}^{k+h} \mathbf{e}_p\left(-cv\right) \right| \left| \sum_{x \in \mathcal{X}} \mathbf{e}_p\left(cf(x)\right) \right| \\ &\quad = O\left(\#X\Delta p^{-1} \sum_{c=1}^{p-1} \left| \sum_{v=k}^{k+h} \mathbf{e}_p\left(cv\right) \right| \right) \\ &\quad = O\left(\#X\Delta \log p\right). \end{aligned}$$

□

More applications a little later . . .

Chapter 3

More Sophisticated Methods

3.1 Extend and Conquer

Here we show that sometimes it is profitable to **extend** our sum over a small set of arbitrary structure to a bigger set (just potentially increasing the size of the sum) with a nice well-studied structure. Certainly we can not do this with the original sum because the terms are complex numbers but this idea can be combined with some tricks. Very often it is used together with the Cauchy inequality in the form

$$\left(\sum_{j=1}^{m} s_j\right)^2 \leq m \sum_{j=1}^{m} s_j^2$$

which holds for any non-negative $s_1, \ldots, s_m$.
We demonstrate this principle on the following very important example. Let $\mathcal{X}$ and $\mathcal{Y}$ be arbitrary subsets of $\mathbb{F}_p$.
Define

$$W_c = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \mathbf{e}_p(cxy)$$

Trivially $|W_c| \leq \#\mathcal{X}\#\mathcal{Y}$. We show that very simple arguments allow us to obtain a bound which is better than trivial for $\#\mathcal{X}\#\mathcal{Y} \geq p$. Thus this bound improves the trivial bound for *very sparse* sets of arbitrary structure!

Theorem 6 *For any sets $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{F}_p$,*

$$|W_c| \leq \left(\#\mathcal{X}\#\mathcal{Y}p\right)^{1/2}.$$

Proof. We have

$$|W_c| = \left|\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \mathbf{e}_p(cxy)\right| \le \sum_{x\in\mathcal{X}}\left|\sum_{y\in\mathcal{Y}} \mathbf{e}_p(cxy)\right|.$$

From the Cauchy inequality,

$$|W_c|^2 \le \#\mathcal{X}\sum_{x\in\mathcal{X}}\left|\sum_{y\in\mathcal{Y}} \mathbf{e}_p(cxy)\right|^2.$$

We **extend** the sums over x to all $x \in \mathbb{F}_p$:

$$|W_c|^2 \le \#\mathcal{X}\sum_{x\in\mathbb{F}_p}\left|\sum_{y\in\mathcal{Y}} \mathbf{e}_p(cxy)\right|^2$$

This is a very important step! We add many more terms to our sums (which we can do because each term is positive). Of course we lose here but our gain is that the sum over x (taken from some mysterious set we have no information about) is now extended to a very nice set.
Now we *Conquer*:

$$\begin{aligned}\sum_{x\in\mathbb{F}_p}\left|\sum_{y\in\mathcal{Y}} \mathbf{e}_p(cxy)\right|^2 &= \sum_{x\in\mathbb{F}_p}\sum_{y_1,y_2\in\mathcal{Y}} \mathbf{e}_p\left(cx\left(y_1-y_2\right)\right)\\ &= \sum_{y_1,y_2\in\mathcal{Y}}\sum_{x\in\mathbb{F}_p} \mathbf{e}_p\left(cx\left(y_1-y_2\right)\right)\\ &= p\sum_{\substack{y_1,y_2\in\mathcal{Y}\\ y_1=y_2}} 1 = \#\mathcal{Y}p.\end{aligned}$$

□

Without any assumptions on $\mathcal{X}$ and $\mathcal{Y}$ this bound remains the best possible.

3.2 Clone, Extend and Conquer

The previous principle works for double sums. Here we show how we can create multiple **clones** of our sum and thus reduce it to a double sum.
As in the previous section we use a very important example to exhibit this principle.

Let g, $\gcd(g,p)=1$, be of multiplicative order t modulo p, that is,

$$g^k \equiv 1 \pmod p \quad \Longrightarrow \quad k \equiv 0 \pmod t.$$

Define

$$S(a,b) = \sum_{x=1}^{t} \mathbf{e}_p\left(ag^x\right)\mathbf{e}_t(bx).$$

The term $\mathbf{e}_t(bx)$ is rather unattractive (and unnatural) but we will see soon why it is needed for some applications, see Theorem 8.
Trivially, $|S(a,b)| \le t$.

Theorem 7 *For any a, b with $\gcd(a,p)=1$,*

$$|S(a,b)| \le p^{1/2}.$$

Proof. The function $\mathbf{e}_p\left(ag^x\right)\mathbf{e}_t(bx)$ is periodic with period t. Thus, for $y = 1,\ldots,t$,

$$\begin{aligned}
S(a,b) &= \sum_{x=1}^{t} \mathbf{e}_p\left(ag^{x+y}\right)\mathbf{e}_t(b(x+y)) \\
&= \mathbf{e}_t(by)\sum_{x=1}^{t} \mathbf{e}_p\left(ag^y g^x\right)\mathbf{e}_t(bx) \\
&= \mathbf{e}_t(by) S(ag^y, b).
\end{aligned}$$

Therefore, we can *clone*:

$$|S(a,b)| = |S(ag^y,b)|.$$

Now we *extend*:

$$t|S(a,b)|^2 = \sum_{y=1}^{t} |S(ag^y,b)|^2 \le \sum_{c=0}^{p-1} |S(c,b)|^2.$$

Finally, we *conquer*:

$$\begin{aligned}
t|S(a,b)|^2 &\le \sum_{c=0}^{p-1} |S(c,b)|^2 \\
&= \sum_{x_1,x_2=1}^{t} \mathbf{e}_t(b(x_1-x_2)) \sum_{c=0}^{p-1} \mathbf{e}_p\left(c\left(g^{x_1}-g^{x_2}\right)\right) \\
&= tp
\end{aligned}$$

because

$$g^{x_1} - g^{x_2} \equiv 0 \pmod p$$

iff

$$x_1 \equiv x_2 \pmod t.$$

□

For some values of t this bound remains the best possible, see also Theorems 10.

3.3 Mordell's Bound

We are now ready to prove something more complicated and less straightforward than our previous estimates.
For a polynomial $f \in \mathbb{F}_p[X]$ of degree $\deg f = n$ we define

$$S(f) = \sum_{x=0}^{p-1} \mathbf{e}_p(f(x)).$$

Without loss of generality we can assume that $f(0) = 0$.
Mordell's method follows the following 3 main stages
Stage I. Cloning: For $\lambda \in \mathbb{F}_p^*$, $\mu \in \mathbb{F}_p$, define

$$f_{\lambda,\mu}(x) = f(\lambda x + \mu) - f(\mu).$$

Obviously $S(f) = S(f_{\lambda,\mu})$ (because $x \to \lambda x + \mu$ is a permutation on $\mathbb{F}_p$).

Stage II. Extending: The leading coefficient of $f_{\lambda,\mu}$ is $A\lambda^n$ where $A \neq 0$ is the leading coefficient of f. There are at least $p(p-1)/n$ distinct polynomials $f_{\lambda,\mu}$:

$$\frac{p(p-1)}{n}|S(f)|^{2n} \leq \sum_{\substack{\deg g \leq n \\ g(0)=0}} |S(g)|^{2n}.$$

Stage III. Conquering: Finally we obtain

$$\begin{aligned}
&\sum_{\substack{\deg g\le n\\ g(0)=0}} |S(g)|^{2n}\\
&\quad= \sum_{\substack{\deg g\le n\\ g(0)=0}} S(g)^n\overline{S(g)}^n = \sum_{\substack{\deg g\le n\\ g(0)=0}} S(g)^n S(-g)^n\\
&\quad= \sum_{\substack{\deg g\le n\\ g(0)=0}} \sum_{x_1,\dots,x_{2n}=0}^{p-1} \mathbf{e}_p\left(\sum_{\nu=1}^{n} g(x_\nu) - \sum_{\nu=n+1}^{2n} g(x_\nu)\right)\\
&\quad= \sum_{x_1,\dots,x_{2n}=0}^{p-1}\\
&\qquad\times \sum_{a_1,\dots,a_n=0}^{p-1} \mathbf{e}_p\left(\sum_{j=1}^{n} a_j\left(\sum_{\nu=1}^{n} x_\nu^j - \sum_{\nu=n+1}^{2n} x_\nu^j\right)\right)\\
&\quad= \sum_{x_1,\dots,x_{2n}=0}^{p-1} \prod_{j=1}^{n}\sum_{a_j=0}^{p-1} \mathbf{e}_p\left(a_j\left(\sum_{\nu=1}^{n} x_\nu^j - \sum_{\nu=n+1}^{2n} x_\nu^j\right)\right)\\
&\quad= p^n T,
\end{aligned}$$

where T is the number of solutions of

$$\sum_{\nu=1}^{n} x_\nu^j \equiv \sum_{\nu=n+1}^{2n} x_\nu^j \pmod p, \qquad j = 1,\dots,n,$$

where $0 \le x_1,\dots,x_{2n} \le p-1$.

The first n symmetric functions of $x_1,\dots,x_n$ and $x_{n+1},\dots,x_{2n}$ are the same. Recalling the Newton formulas we see that they are roots of the same polynomial of degree n. Therefore they are permutations of each other.

There are p^n values for $x_1,\dots,x_n$ and for each fixed values of $x_1,\dots,x_n$ there are at most $n!$ values for the other n variables $x_{n+1},\dots,x_{2n}$. Therefore

$$T \le n!p^n.$$

This yields

$$|S(f)| \le c(n)p^{1-1/n}$$

where $c(n) = (n\,n!)^{1/2n} \approx (n/e)^{1/2}$.

3.4 Shorter Sums ... but Large Bound

Here we show a general principle how the problem of bounding incomplete sums to the problem of bounding *almost the same* complete sums. Unfortunately, we lose a little bit, the bound because bigger by a logarithmic factor. For g, $\gcd(g,p)=1$, of multiplicative order t modulo p, define **incomplete** sums

$$T(a;N) = \sum_{x=1}^{N} \mathbf{e}_p\left(ag^x\right).$$

Theorem 8 *For any a with $\gcd(a,p)=1$ and $N \le t$*

$$|T(a;N)| = O(p^{1/2}\log p).$$

Proof. We have

$$\begin{aligned}
|T(a;N)| &= \left|\sum_{x=1}^{t} \mathbf{e}_p\left(ag^x\right) \frac{1}{t}\sum_{b=0}^{t-1}\sum_{y=1}^{N} \mathbf{e}_t(b(x-y))\right| \\
&= \frac{1}{t}\left|\sum_{b=0}^{t-1} S(a,b) \sum_{y=1}^{N} \mathbf{e}_t(-by)\right| \\
&\le \frac{1}{t}\sum_{b=0}^{t-1} |S(a,b)| \left|\sum_{y=1}^{N} \mathbf{e}_t(-by)\right| \\
&\le \frac{p^{1/2}}{t}\sum_{b=0}^{t-1} \left|\sum_{y=1}^{N} \mathbf{e}_t(-by)\right| = O(p^{1/2}\log p)
\end{aligned}$$

by Theorem 7 and Lemma 4. □

Chapter 4

Some Strongest Known Results

4.1 Weil's Kingdom

Using algebraic geometry tools due to Andre Weil [79] (an upper bound for the number of solutions of equations $F(x,y) = 0$ in finite fields) one can prove much stronger bounds for various sums with

- polynomials;
- rational functions;
- algebraic functions.

Here we present only one of such bounds in the following form given by C. Moreno and O. Moreno

Theorem 9 *For any polynomials $g(X), h(X) \in \mathbb{F}_p[X]$ such that the rational function $f(X) = h(X)/g(X)$ is not constant on $\mathbb{F}_p$, the bound*

$$\left| \sum_{\substack{x \in \mathbb{F}_p \\ g(x) \neq 0}} \mathbf{e}_p\left(f(x)\right) \right| \leq \left(\max\{\deg g\,, \deg h\} + r - 2\right) p^{1/2} + \delta$$

holds, where

$$(r, \delta) = \begin{cases} (v, 1), & \text{if } \deg h \leq \deg g, \\ (v+1, 0), & \text{if } \deg h > \deg g, \end{cases}$$

and v is the number of distinct zeros of $g(X)$ in the algebraic closure of $\mathbb{F}_p$.

In the special case when $f(X)$ is a not constant polynomial of degree $\deg f = n$ the bound takes its well-known form

$$\left|\sum_{x\in\mathbb{F}_p} \mathbf{e}_p\left(f(x)\right)\right| \le (n-1)\,p^{1/2}. \tag{4.1}$$

Nowadays we have a pure elementary alternative to the algebraic geometry which is due to S. A. Stepanov, N. M. Korobov, H. Stark, W. Schmidt and to several other researchers.

Surprisingly enough, in some special cases elementary method gives much stronger results. Such improvements are due to A. Garcia and F. Voloch, D. Mit'kin, R. Heath-Brown and S. V. Konyagin, for more details see [25].

It is important to remember that

$$\boxed{\text{“elementary”} \neq \text{“simple”}}$$

"Elementary" merely means that there is no explicit use of any algebraic geometry notions and tools.

For multivariate polynomials an analogue of (4.1) is due to P. Deligne [15] but it requires some special conditions on the polynomial in the exponent which are not so easy to verify. This limits the range of applications of that bound, while the Weil bound (4.1) is very easy to apply.

4.2 Exponential Functions

Exponential functions form another natural family of functions which arise in many applications. The problem of estimating exponential sums with exponential functions has a long history, we refer to [38, 39, 40, 45, 58, 59, 65] for more details.

Using some improvements of the Weil bound due to R. Heath-Brown and S. V. Konyagin [25], one can improve Theorem 7. Namely the following result has been obtained by S. V. Konyagin and I. E. Shparlinski [38], Theorem 3.4.

Theorem 10 *For any a, b with $\gcd(a,p) = 1$,*

$$|S(a,b)| \le \begin{cases} p^{1/2}, & \text{if } t \ge p^{2/3}; \\ p^{1/4}t^{3/8}, & \text{if } p^{2/3} > t \ge p^{1/2}; \\ p^{1/8}t^{5/8}, & \text{if } p^{1/2} > t \ge p^{1/3}; \end{cases}$$

holds.

The main challenge is to obtain nontrivial bounds for as small values of t as possible. Theorem 10 works only for $t \geq p^{1/3+\varepsilon}$. For almost all primes Theorem 5.5 of [38] provides a nontrivial bound for $t \geq p^{\varepsilon}$. We present it in the form given in [53].

Theorem 11 *Let Q be a sufficiently large integer. For any $\varepsilon > 0$ there exists $\delta > 0$ such that for all primes $p \in [Q, 2Q]$, except at most $Q^{5/6+\varepsilon}$ of them, and any element $g_{p,T} \in \mathbb{F}_p$ of multiplicative order $T \geq p^{\varepsilon}$ the bound*

$$\max_{\gcd(c,p)=1} \left| \sum_{x=0}^{T-1} \mathbf{e}_p\left(c g_{p,T}^x\right) \right| \leq T^{1-\delta}$$

holds.

4.3 More Applications

Combining the Weil bound 4.1 and Theorem 5 we obtain that for any polynomial f of degree n

$$\max_k \max_{0 \leq h \leq p-1} |N_f(k,h) - h| = O(np^{1/2} \log p). \tag{4.2}$$

We recall that an number $a \not\equiv 0 \pmod p$ is called a *quadratic residue* if the congruence $a \not\equiv x^2 \pmod p$ has a solution and is is called a *quadratic non-residue* otherwise. Numbers a with $a \equiv 0 \pmod p$ are not clasified.

Using (4.2) for the quadratic polynomial $f(x) = x^2$ we see in any interval $[k, k+h-1]$ the imbalance between the number of quadratic residues modulo p and non-residues is at most $O(p^{1/2} \log p)$. This is the famous *Polya–Vinogradov* inequality.

More precisely, let us denote by $Q_+(k,h)$ and $Q_-(k,h)$ the numbers of quadratic residues and non-residues, respectively, in the interval $[k, k+h-1]$.

Theorem 12 *The bound*

$$\max_k \max_{0 \leq h \leq p-1} \left| Q_\pm(k,h) - \frac{h}{2} \right| = O(p^{1/2} \log p)$$

holds.

Proof. Because the residue ring modulo p is a field we see that if $a \not\equiv 0 \pmod p$ and the congruence $a \not\equiv x^2 \pmod p$ has a solution, then it has two distinct solutions. Taking into account that an interval $[k, k+h-1]$ with $0 \le h \le p-1$ contains at most one zero, we obtain the inequalities

$$\frac{1}{2}N_f(k,h) - 1 \le Q_+(k,h) \le \frac{1}{2}N_f(k,h)$$

and

$$h - 1 \le Q_+(k,h) + Q_-(k,h) \le h.$$

Using (4.2) we obtain the desired result. $\square$

Similarly, Theorems 10 and Theorems 11 can be used to study the distribution of the values of g^x in short intervals, see [38, 65, 66] for numerous applications of this type of result to cryptography, coding theory and computer science.

4.4 What Else Can We Estimate?

There are several other classes of exponential sums which have attracted much of attention of experts in analytical number theory. Here we present a short outline of such classes.

- Exponential sums with composite denominator

  $$S(f) = \sum_{x=0}^{p-1} \mathbf{e}_q(f(x)),$$

  where $q \ge 1$ is an integer, $f \in \mathbb{Z}[X]$. These sums are very well studied, thanks to works of Hua Loo Keng, Vasili Nechaev, Sergei Stečkin, see [32, 33, 72].

- Exponential sums with recurring sequences For linear recurring sequences such estimates are due to N. M. Korobov and H. Niederreiter, see [45, 40, 58, 59, 65]. For nonlinear recurring sequences such estimates are due to H. Niederreiter and I. E. Shparlinski, see [60].

- *H. Weyl, P. van der Corput, I. M Vinogradov, N. .M. Korobov*: sums with polynomials with irrational coefficients ... *not much progress since 1947.*

- It is easy to see that $\mathbf{e}_p(\cdot)$ is an additive character of $\mathbb{F}_p$. Similar results are know for additive and multiplicative characters of arbitrary finite fields and residue rings. Although usually for sums of multiplicative characters the theory follows the same path as for exponential sums there are some exceptions. For example, there is no analogue of Theorem 4 for multiplicative character sums. On the other hand, the celebrated Burgess bound [12] has no analogue for exponential sums.

- Thousands of less general results for various interesting (and not so) special cases.

Chapter 5

Applications to Coding Theory

5.1 Direct Applications

Many coding theory questions can immediately be formulated as questions about bound of exponential sums:

- correlation and autocorrelation, see [3, 4, 5, 23, 17, 26, 27, 28, 29];
- Minimal distance of BCH codes [47];
- Size of Varshamov–Mazur codes for asymmetric channels [38, 48, 65].

Surprisingly enough, it works the other way as well. Some coding theory lower bounds can be applied to obtain very tight lower bounds for exponential sums [6, 44, 61, 63, 73, 74]. One can certainly argue about the importance of lower bounds because all known applications are based on upper bounds. Nevertheless they certainly improve our understanding of the area and are an intrinsic part of the theory of exponential sums.
Several other interrelations between exponential sums and coding theory, which enrich both areas, can be found in [65].

5.2 Less Obvious Applications: Dimension of BCH Codes

Let q be a prime power and let n be an integer with $\gcd(n, q) = 1$.
Denote by t the multiplicative order of q modulo n; and fix an element $\alpha \in \mathbb{F}_{q^t}^*$ of multiplicative order n (it exists because $n \mid q^t - 1$);.

Let l be an integer. To construct a BCH code with *constructive distance* Δ we consider the polynomial g over $\mathbb{F}_q$ of the smallest degree such that

$$g\left(\alpha^{l+y}\right) = 0, \qquad y = 1, \dots, \Delta - 1,$$

and consider the cyclic code of length n with g as the generator polynomial. That is the linear space of dimension $k = n - \deg g$ of n-dimensional vectors $(a_0, \dots, a_{n-1}) \in \mathbb{F}_q^n$ such that

$$a_0 + a_1 Z + \dots + a_{n-1} Z^{n-1} \equiv 0 \pmod{g(Z)}.$$

Generally for every code there are three parameters of interest: the length, the minimal distance and the dimension. For a BCH code the length n is given, the minimal distance d is at least the constructive distance Δ (and this bound is known to be tight in many cases [47]). The question about the dimension is more interesting. Of course, $t \le \deg g \le Dt$, thus the dimension $n - t \ge k \ge n - (\Delta - 1)t$. To get something stronger one should study the structure of the roots of g in more detail.
First of all we make an observation that all roots of g are powers of α because trivially

$$g(Z) \mid \prod_{y=1}^{\Delta-1} \prod_{x=1}^{t} \left(Z - \alpha^{(l+y)q^x}\right).$$

We also remark that α^j is a root of g is and only if

$$jq^x \equiv l + y \pmod{n},$$

for some $x = 1, \dots, t$ and $y = 1, \dots, \Delta - 1$.
The code is the linear space of dimension $k = n - \deg g$ of n-dimensional vectors $(a_0, \dots, a_{n-1}) \in \mathbb{F}_q^n$ such that

$$a_0 + a_1 Z + \dots + a_{n-1} Z^{n-1} \equiv 0 \pmod{g(Z)}.$$

We have

$$\Delta - 1 \le \deg g \le (\Delta - 1)t$$

and

$$n - \Delta + 1 \ge k \ge n - (\Delta - 1)t.$$

To improve one should study g in more detail.
Observations:

- all roots of g are powers of α because

$$g(Z) \mid \prod_{y=1}^{\Delta-1} \prod_{x=1}^{t} \left(Z - \alpha^{(l+y)q^x}\right);$$

- α^j is a root of g iff $jq^x \equiv l + y \pmod{n}$, for some $x = 1, \dots, t$ and $y = 1, \dots, \Delta - 1$.

$J(q, n, \Delta)$ = the maximal dimension of q-ary generalized BCH codes of length n and of designed distance Δ taken over all $l = 0, \dots, n-1$.
$J(q, n, \Delta)$ = the number of $j = 0, 1, \dots, n-1$ for which the congruence

$$jq^x \equiv l + y \pmod{n}, \qquad 1 \le x \le t,\ 1 \le y \le \Delta - 1,$$

is not solvable.
Thus the original questions has been reduced to a question about the distribution of values of an exponential function to which our technique can be applied.
For a divisor d of n denote by t_d the multiplicative order q modulo d (thus $t = t_n$).

Lemma 13 *For any $d \mid n$, the bound $t_{n/d} \ge t/d$ holds.*

Lemma 14 *For any integer a, b, the congruence*

$$aq^x \equiv bq^y \pmod{n}, \qquad 1 \le x, y \le t$$

is solvable only when $\gcd(a, n) = \gcd(b, n) = d$, and in this case for the number of solutions $N(a, b)$ the bound

$$N(a, b) \le td$$

holds.

Proof. As $\gcd(q, n) = 1$, the condition on a and b is evident. Also it is evident that for any fixed x there are at most $t/t_{n/d}$ possible values for y, hence $N(a, b) \le t^2/t_{n/d} \le td$ because of Lemma 13. □
We define the sums

$$T(a, h) = \sum_{u=1}^{h} \mathbf{e}(au/n), \quad W_d(h) = \sum_{\gcd(a,n)=d} |T(a, h)|^2,$$

where d is a divisor of n, $d \mid n$.

Lemma 15 *For any $d \mid n$ with $d < n$, the bound*

$$W_d(h) \le nh/d$$

holds.

Proof. Denote $m = n/d$. We have

$$W_d(h) \le \sum_{a=0}^{n/d-1} |T(ad,h)|^2 - h^2 = mM - h^2,$$

where M is the number of solutions of the congruence

$$u \equiv v \pmod m, \qquad 1 \le u, v \le h.$$

Write $h = km + r$ with $0 \le r \le m-1$, then $M = r(k+1)^2 + (m-r)k^2 = k^2m + 2kr + r$. Therefore

$$\begin{aligned} W_d(h) &\le mM - h^2 = k^2m^2 + 2kmr + mr - h^2 \\ &= (h-r)^2 + 2r(h-r) + mr - h^2 = r(m-r) \\ &\le rm \le hm = nh/d. \end{aligned}$$

□

Theorem 16 *The bound*

$$J(q,n,\Delta) \le \frac{4n^3}{(\Delta-1)^2 t}.$$

holds.

Proof. Let $h = \lfloor \Delta/2 \rfloor$ and let N_j denote the number of solutions of the congruence

$$jq^x \equiv l + h + u - v \pmod n,$$

$$x = 1, \ldots, t, \quad u, v = 1, \ldots, h.$$

Then $J(q,n,\Delta) \le |\mathbf{I}(q,n,\Delta)|$ where $\mathbf{I}(q,n,\Delta)$ is the set of $j = 0, 1, \ldots, n-1$ for which this congruence is unsolvable, that is, $N_j = 0$.
Set

$$S(a) = \sum_{x=1}^{t} \mathbf{e}(aq^x/n).$$

Then $N_j = th^2/n + R_j/n$ where

$$R_j = \sum_{a=1}^{n-1} S(aj)|T(a,h)|^2 \mathbf{e}(-a(l+h)/p).$$

Let us consider

$$R = \sum_{j=0}^{n-1} R_j^2.$$

We have

$$\begin{aligned} R &= \sum_{j=0}^{n-1} \sum_{a,b=1}^{n-1} S(aj)S(bj) \\ &\qquad \times |T(a,h)|^2 |T(b,h)|^2 \mathbf{e}\left(-(a+b)(l+h)/p\right) \\ &= \sum_{a,b=1}^{n-1} |T(a,h)|^2 |T(b,h)|^2 \mathbf{e}\left(-(a+b)(l+h)/p\right) \\ &\qquad \times \sum_{j=0}^{n} S(aj)S(bj). \end{aligned}$$

Then,

$$\begin{aligned} \sum_{j=0}^{n-1} S(aj)S(bj) &= \sum_{x,y=1}^{t} \sum_{j=0}^{n-1} \mathbf{e}\left(j(aq^x + bq^y)/n\right) \\ &= nN(a,-b). \end{aligned}$$

For all divisors $d \mid n$ we gather together all terms corresponding to a and b with

$$\gcd(a,n) = \gcd(b,n) = d.$$

Applying Lemma 14, we obtain

$$\begin{aligned} R &= n \sum_{d \mid n, d<n} \; \sum_{\gcd(a,n)=\gcd(b,n)=d} |T(a,h)|^2 |T(b,h)|^2 \\ &\qquad \times N(a,-b)\mathbf{e}\left((a+b)(l+h)/p\right) \\ &\le nt \sum_{d \mid n, d<n} dW_d(h)^2 \\ &\le nt \max_{d \mid n, d<n} dW_d(h) \sum_{d \mid n, d<n} W_d(h). \end{aligned}$$

From Lemma 15 and the identity

$$\sum_{\substack{d \mid n \\ d<n}} W_d(h) = \sum_{a=1}^{n-1} |T(a,h)|^2 = nh - h^2$$

we derive

$$R \leq n^3 h^2 t.$$

Since $R_j = -h^2 t$ for $j \in \mathbf{I}(q,n,\Delta)$ then

$$|\mathbf{I}(q,n,\Delta)| h^4 t^2 = \sum_{j=0}^{n-1} R_j^2 \leq n^3 h^2 t.$$

Taking into account that $h \geq (\Delta - 1)/2$, we obtain the result. □

It is useful to keep in mind that *exponential sums do not always win.* For certain values of parameters the following elementary statement provides a sharper bound.

Theorem 17 *The bound*

$$J(q,n,\Delta) \leq 2e^{1/2} n^{1-\alpha_q(\delta)}$$

holds, where $\delta = (\Delta - 1)/n$ *and*

$$\alpha_q(\delta) = \frac{\delta}{2\ln(3q/\delta)}.$$

Thus this can be taken as an encouragement to study other number theoretic techniques.

Chapter 6

Applications to Cryptography

6.1 Distribution of Some Cryptographic Primitives

6.1.1 Security of Exponentiation with Precomputation

Let g be an element of order t modulo a prime number p, that is

$$g^T \equiv 1 \pmod p \quad \Leftrightarrow \quad t|T.$$

Let r be the bit length of t, $2^{r-1} \le t \le 2^r - 1$.
Many signature schemes use exponentiation $g^x \pmod p$ for a "random" x.
Using repeated squaring this takes about $1.5r$ multiplications on average and about $2r$ operations in the worst case.
One of the possible ways to speed-up exponentiation is to precompute the values $g^{2^j} \pmod p$, $j = 0, \dots, r$. Then computing $g^x \pmod p$ takes $0.5r$ multiplications on average, r multiplications in the worst case.
Main Problem: How can we generate *secure* pairs (x, g^x) faster (for some special x)?
Secure: Finding x from g^x for the values of x generated by this method should be as hard as for a random $x \in [0, M-1]$.
In 1998, V. Boyko, M. Peinado and R. Venkatesan [10] proposed the following algorithm
Given $n \ge k \ge 1$:

Preprocessing Step: Generate n random integers $\alpha_1, \dots, \alpha_n \in \mathbb{Z}_M$. Compute $\beta_j \equiv g^{\alpha_j} \pmod p$ and store the values of α_j and β_j in a table, $j = 1, \dots, n$.

Pair Generation: Generate a random set $S \subseteq \{1, \dots, n\}$ of cardinality $\#S = k$. Compute

$$x \equiv \sum_{j \in S} \alpha_j \pmod M, \quad X \equiv \prod_{j \in S} \beta_j \equiv g^b \pmod p.$$

Cost: $k - 1$ modular multiplications.

It is easy to see that

$$X \equiv g^x \pmod p.$$

In 1999, P. Q. Nguyen, I. E. Shparlinski and J. Stern [55] proposed the following generalization of this scheme which involved one more integer parameter h.

Given $n \geq k \geq 1$ and $h \geq 2$:

Preprocessing Step: Generate n random integers $\alpha_1, \dots, \alpha_n \in \mathbb{Z}_M$. Compute $\beta_j \equiv g^{\alpha_j} \pmod p$ and store the values of α_j and β_j in a table, $j = 1, \dots, n$.

Extended Pair Generation: Generate a random set $S \subseteq \{1, \dots, n\}$, $|S| = k$ and for each $j \in S$ select a random integer $x_j \in \{0, \dots, h-1\}$. Compute

$$x \equiv \sum_{j \in S} \alpha_j x_j \pmod M, \quad X \equiv \prod_{j \in S} \beta_j^{x_j} \pmod p.$$

Cost: $k + h - 3$ modular multiplications.

One verifies that the congruence

$$X \equiv g^x \pmod p$$

holds. The cost estimate (which is better than naive $O(k \log h)$) follows from a result of [11].

Finally, using some bounds of exponential sums and to establish some results about the uniformity of distribution of sums

$$\sum_{j \in S} \alpha_j x_j \pmod M, \quad x_j \in \{1, \dots, h-1\},$$

the security of this scheme was proved in [55]. We present this result in an informal way and refer to [55] for exact formulations (which formalises the notion of security).

Theorem (informally). *Let $n = \gamma r$ with some $\gamma > 0$ there are values of k and h with*

$$k + h = O(r/\log r)$$

and such that the scheme is as secure as the generator $x \to g^x \pmod p$ for arbitrary x.

The most important characteristics of this scheme are

Table size: linear in r (say $n = r$)

Speed-up: $\log r \to \infty$.

6.1.2 Diffie-Hellman Triples and RSA Pairs

Let g be a primitive element modulo p.

The following assumption is known as the *Diffie–Hellman Indistinguishability Assumption*: It is feasible to distinguish between *Diffie–Hellman triples* (g^x, g^y, g^{xy}) with random x and y and random triples $(u, v, w) \in \mathbb{F}_p^3$?

One of the possible (very naive approaches) to disprove this assumption would be to find some statistically "visible" singularities in differences in the distribution of the triples

$$(g^x, g^y, g^{xy}), \qquad x, y = 1, \ldots, p-1.$$

However, R. Canetti, J. Friedlander and I. E. Shparlinski [14] in 1997, and a stronger form, R. Canetti, J. B. Friedlander, S. V. Konyagin, M. Larsen, D. Lieman and I. Shparlinski [13] in 1999, proved that Diffie-Hellman triples are uniformly distributed.

Recently, J. B. Friedlander and I. E. Shparlinski [18] obtained similar statement for Diffie-Hellman triples with "sparse" x and y (one can use such x and y in order to speed-up computation). It is shown in [18] that Diffie-Hellman triples with x and y having at most $0.35 \log p$ nonzero digits are uniformly distributed.

Surprisingly enough, several results of [13] play a central role in studying a related problem about the distribution of *RSA pairs* (x, x^e) in the residue ring modulo m, see [69].

6.2 Lattices and Exponential Sums

6.2.1 Introduction and Notation

In this section we describe how a rather unusual combination of two celebrated number theoretic techniques, namely, bounds of *exponential sums* and *lattice reduction* algorithms, provides a powerful cryptographic tool. It can be applied to both proving several security results and designing new attacks.

For example, it has been used to prove certain bit security results for the Diffie-Hellman key exchange system, for the Shamir message passing scheme and for the XTR cryptosystem. It has also been used to design provable attacks on the Digital Signature Scheme and its modifications, including the Nyberg–Rueppel scheme, which are provably insecure under certain conditions.

Here we explain how these two techniques get together, outline several important applications applications and discuss some open problems on exponential sums which arise in this context and which need to be solved before any further progress in this area can be achieved.

Let p denote a prime number and let $\mathbb{F}_p$ denote the finite field of p elements. For integers s and $m \geq 1$ we denote by $\lfloor s \rfloor_m$ the remainder of s on division by m. For a prime p and $\ell > 0$ we denote by $\mathrm{MSB}_{\ell,p}(x)$ any integer u such that

$$|\lfloor x \rfloor_p - u| \leq p/2^{\ell+1}. \tag{6.1}$$

Roughly speaking, $\mathrm{MSB}_{\ell,p}(x)$ gives ℓ most significant bits of x however this definition is more flexible and suits better our purposes. In particular we remark that ℓ in the inequality (6.1) need not be an integer.

Throughout this paper $\log z$ denotes the binary logarithm of $z > 0$.

The implied constants in symbols 'O' may occasionally, where obvious, depend on the small positive parameters ε and are absolute otherwise.

6.2.2 Hidden Number Problem and Lattices

We start with a certain algorithmic problem, introduced in 1996 by Boneh and Venkatesan [8, 9], which seemingly has nothing in common with exponential sums. Namely we consider the following

> HIDDEN NUMBER PROBLEM, **HNP**: *Recover a number $\alpha \in \mathbb{F}_p$ such that for many known random $t \in \mathbb{F}_p^*$ we are given $\mathrm{MSB}_{\ell,p}(\alpha t)$ for some $\ell > 0$.*

The paper [8] also contains a polynomial time algorithm to solve this problem (with ℓ of order $\log^{1/2} p$). The most important ingredient of this algorithm is lattice reduction.

We briefly review a few results and definitions. For general references on lattice theory and its important cryptographic applications, we refer to the recent surveys [56, 57].

Let $\{\mathbf{b}_1, \dots, \mathbf{b}_s\}$ be a set of linearly independent vectors in $\mathbb{R}^s$. The set of vectors

$$L = \left\{ \sum_{i=1}^{s} n_i \mathbf{b}_i \mid n_i \in \mathbb{Z} \right\},$$

is called an s-dimensional full rank lattice. The set $\{\mathbf{b}_1, \dots, \mathbf{b}_s\}$ is called a *basis* of L, and L is said to be spanned by $\{\mathbf{b}_1, \dots, \mathbf{b}_s\}$.

One of the most fundamental problems in this area is the *closest vector problem*, **CVP**: given a basis of a lattice L in $\mathbb{R}^s$ and a target vector $\mathbf{u} \in \mathbb{R}^s$, find a lattice vector $\mathbf{v} \in L$ which minimizes the Euclidean norm $\|\mathbf{u} - \mathbf{v}\|$ among all lattice vectors. It is well know that **CVP** is **NP**-hard (see [56, 57] for references). However, its approximate version [2] admits a polynomial time algorithm which goes back to the lattice basis reduction algorithm of Lenstra, Lenstra and Lovász [41].

It has been remarked in Section 2.1 of [49] and then in Section 2.4 of [56] and Section 2.4 of [57] that the following statement holds which is somewhat stronger than that usually used in the literature.

Theorem 18 *There exists a polynomial time algorithm which, given an s-dimensional full rank lattice L and a vector $\mathbf{r} \in \mathbb{R}^s$, finds a lattice vector $\mathbf{v}$ satisfying the inequality*

$$\|\mathbf{v} - \mathbf{r}\| \le 2^{O\left(s \log^2 \log s / \log s\right)} \min \left\{ \|\mathbf{z} - \mathbf{r}\|, \quad \mathbf{z} \in L \right\}.$$

Proof. The statement is a combination of Schnorr's modification [64] of the lattice basis reduction algorithm of Lenstra, Lenstra and Lovász [41] with a result of Kannan [34] about reduction of the **CVP** to the approximate shortest vector problem. □

One can also use a probabilistic analogue [1] of Theorem 18 which gives a slightly better constant.

We are now prepared to sketch the main ideas of [8] to solve the **HNP**. Let $d \ge 1$ be integer. Given t_i, $u_i = \mathrm{MSB}_{\ell,p}(\alpha t_i)$, $i = 1, \dots, d$, we build the lattice $\mathcal{L}(p, \ell, t_1, \ldots, t_d)$ spanned by the rows of the matrix:

$$\begin{pmatrix} p & 0 & \ldots & 0 & 0 \\ 0 & p & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & \ldots & p & 0 \\ t_1 & t_2 & \ldots & t_d & 1/2^{\ell+1} \end{pmatrix}.$$

and notice

$$\mathbf{w} = (\lfloor \alpha t_1 \rfloor_p, \ldots, \lfloor \alpha t_d \rfloor_p, \alpha/2^{\ell+1}) \in \mathcal{L}(p, \ell, t_1, \ldots, t_d).$$

This vector is very close to the *known* vector $\mathbf{u} = (u_1, \ldots, u_d, 0)$ (at the distance of order $p2^{-\ell}$). Thus applying one of the lattice reduction algorithms one can *hope* to recover $\mathbf{v}$ and thus α. In order to make this algorithm *rigorous* one needs to show that (for almost all choices of $t_1, \ldots, t_d \in \mathbb{F}_p$ there is no other lattice vector which is close to $\mathbf{u}$. Namely, taking into account the "stretching" factor in the algorithm of Lemma 18, we have to show that there are very few d-tuples $(t_1, \ldots, t_d) \in \mathbb{F}_p^d$ for which the lattice $\mathcal{L}(p, \ell, t_1, \ldots, t_d)$ has a vector $\mathbf{v} \neq \mathbf{w}$ and such that

$$\|\mathbf{v} - \mathbf{u}\| \le p2^{-\ell} \exp\left(O\left(\frac{d \log^2 \log d}{\log d}\right)\right).$$

The last inequality implies that

$$\|\mathbf{v} - \mathbf{w}\| \le p2^{-\ell} \exp\left(O\left(\frac{d \log^2 \log d}{\log d}\right)\right) \tag{6.2}$$

which is our main tool.
Any vector $\mathbf{v} \in \mathcal{L}(p, \ell, t_1, \ldots, t_d)$ is of the form

$$\mathbf{v} = \left(\beta t_1 - \lambda_1 p, \ldots, \beta t_d - \lambda_d p, \beta/2^{\ell+1}\right),$$

with some integers β and $\lambda_1, \ldots, \lambda_d$. Thus (6.2) implies that for all $i = 1, \ldots, d$ we have

$$(\alpha - \beta) t_i \equiv y_i \pmod{p} \tag{6.3}$$

for some $y_i \in [-h, h]$ where

$$h = p2^{-\ell} \exp\left(O\left(\frac{d \log^2 \log d}{\log d}\right)\right).$$

The probability

$$\Pr_{y\in\mathbb{F}_p}[\gamma t = y \pmod p) \mid y \in [-h,h]] \le \frac{2h+1}{p} \tag{6.4}$$

for any $\gamma \neq 0$.

Therefore the probability P that the condition (6.3) holds for all $i = 1, \ldots, d$ and at least one $\beta \neq \alpha$, is at most

$$P \le (p-1)\left(\frac{2h+1}{p}\right)^d \le p(3h/p)^d = p2^{-\ell d}\exp\left(O\left(\frac{d^2\log^2\log d}{\log d}\right)\right).$$

Thus if

$$\ell = \left\lceil C\frac{\log^{1/2} p \log\log\log p}{\log\log p}\right\rceil \quad \text{and} \quad d = 2\left\lceil\frac{\log p}{\ell}\right\rceil$$

with some absolute constant $C > 0$ then the lattice reduction algorithm returns $\mathbf{v}$ with probability exponentially close to 1.

6.2.3 Extended Hidden Number Problem, Lattices and Exponential Sums

It has turned out that for many applications, including some results about the bit security of Diffie-Hellman, Shamir and some other cryptosystems [21, 22, 67, 68, 70] and rigorous results on attacks (following the heuristic arguments of [31, 52]) on the DSA and DSA-like signature schemes [16, 53, 54], the condition that t is selected uniformly at random from $\mathbb{F}_p$ is too restrictive.

It has been systematically exploited in [16, 21, 22, 53, 54, 67, 68, 70] that the method of [8] can be extended to the case where t is selected from a sequence $\mathcal{T}$ having some uniformity of distribution property.

Accordingly, we consider the following:

> Extended Hidden Number Problem, **EHNP**: *Recover a number $\alpha \in \mathbb{F}_p$ such that for many known random $t \in \mathcal{T}$ we are given* $\mathrm{MSB}_{\ell,p}(\alpha t)$ *for some $\ell > 0$.*

If $\mathcal{T} = \mathbb{F}_p$ then rather simple counting arguments of Section 6.2.2 show that the number of d-tuples $(t_1, \ldots, t_d) \in \mathbb{F}_p^d$ for which the algorithm of Lemma 18 returns a false vector is exponentially small. However for other sequences $\mathcal{T}$ one needs a result about the uniformity of distribution of $\mathcal{T}$.

In the quantitative form which is based on best known lattice reduction algorithms [1, 2, 34, 35, 41, 56, 57, 64] this has been obtained in [53].

Recall that the *discrepancy* of an N-element sequence $\Gamma = \{\gamma_1, \dots, \gamma_N\}$ of elements of the interval $[0, 1]$ is defined as

$$\mathcal{D}(\Gamma) = \sup_{J \subseteq [0,1]} \left| \frac{A(J, N)}{N} - |J| \right|,$$

where the supremum is extended over all subintervals J of $[0, 1]$, $|J|$ is the length of J, and $A(J, N)$ denotes the number of points γ_n in J for $0 \le n \le N - 1$.

We say that a finite sequence $\mathcal{T}$ of integers is Δ-*homogeneously distributed modulo* p if for any integer a, with $\gcd(a, p) = 1$ the discrepancy of the sequence $\{\lfloor at \rfloor_p / p\}_{t \in \mathcal{T}}$ is at most Δ.

In this case the arguments of Section 6.2.2 go through with only one change, namely (6.4) becomes

$$\Pr_{y \in \mathcal{T}} [\gamma t = y \pmod p \mid y \in [-h, h]] \le \frac{2h + 1}{p} + \Delta.$$

This leads to the following result from [53] which extends the algorithm of [8] to the **EHNP** with a general sequence $\mathcal{T}$.

Theorem 19 *For a prime p, define $\ell = \lceil \log^{1/2} p \rceil + \lceil \log \log p \rceil$, and $d = 2 \left\lceil \log^{1/2} p \right\rceil$. Let $\mathcal{T}$ be a $2^{-\log^{1/2} p}$-homogeneously distributed modulo p sequence of integer numbers. There exists a deterministic polynomial time algorithm $\mathcal{A}$ such that for any fixed integer α in the interval $[0, p - 1]$, given a prime p and $2d$ integers*

$$t_i \quad \textit{and} \quad u_i = \mathrm{MSB}_{\ell,p}(\alpha t_i), \qquad i = 1, \dots, d,$$

its output satisfies for sufficiently large p

$$\Pr_{t_1, \dots, t_d \in \mathcal{T}} [\mathcal{A}(p, t_1, \dots, t_d; u_1, \dots, u_d) = \alpha] \ge 1 - 2^{-(\log p)^{1/2} \log \log p}$$

if $t_1, \dots, t_d$ are chosen uniformly and independently at random from the elements of $\mathcal{T}$.

It follows from Corollary 3.11 of [59], that $\mathcal{T}$ is Δ-homogeneously distributed modulo p with

$$\Delta = O\left(\frac{\log p}{\#\mathcal{T}} \max_{c=1,\dots,p-1} \left| \sum_{t \in \mathcal{T}} \exp(2\pi i c t / p) \right| \right). \tag{6.5}$$

The proof follows the same lines as the proof of Theorem 5 and is rather standard if one does not care about the hidden constant in the 'O'-symbol. However in order to get a small constant there one has to go through some technical complications, we refer to [59] for more details.

Therefore, in order to apply this result one can establish the uniformity of distribution of various sequences of $\mathcal{T}$ arising in cryptographic applications and thus one needs to estimate *exponential sums* with elements of $\mathcal{T}$. Thus bounds of exponential sums enter the problem. It has turned out that in some cases relevant exponential sums are well studied in number theory, and thus the corresponding cryptographic result follows immediately, for example, see Section 6.2.4. On the other hand, in some case the exponential sums are of very unusual structure which has no meaningful number theoretic interpretations and thus they have required special treatment, for example, see Section 6.2.5.

6.2.4 Bit Security of the Diffie–Hellman Secret Key

We recall the problem which underlies the Diffie–Hellman key exchange system: given an element g of order τ modulo p, find an efficient algorithm to recover Diffie–Hellman secret key $K = \lfloor g^{xy} \rfloor_p$ from $\lfloor g^x \rfloor_p$ and $\lfloor g^y \rfloor_p$.

Typically, either $\tau = p - 1$ (thus g is a primitive root) or $\tau = q$, a large prime divisor of $p - 1$.

The size of p and τ is determined by the present state of art in the *discrete logarithm problem*. Typically, p is at least about 500 bits, τ is at least about 160 bits.

However after the common DH key $K = \lfloor g^{xy} \rfloor_p$ is established, only a small portion of bits of K will be used as a common key for some pre-agreed *private* key cryptosystem.

Thus a natural question arises: *Assume that finding K is infeasible, is it still infeasible to find certain bits of K?*

In 1996, Boneh and Venkatesan [8] found very elegant links between the **EHNP** and the above problem.

Indeed, assume there is an efficient algorithm to find ℓ most significant bits of $\lfloor g^{xy} \rfloor_p$ from from $X = \lfloor g^x \rfloor_p$ and $Y = \lfloor g^y \rfloor_p$. Then, given $A = \lfloor g^a \rfloor_p$ and $B = \lfloor g^b \rfloor_p$ one can select a random $u \in [0, \tau - 1]$ one can apply the above algorithm to A and $U = \lfloor B g^u \rfloor_p$ getting

$$\mathrm{MSB}_{\ell,p}\left(g^{a(b+u)}\right) = \mathrm{MSB}_{\ell,p}\left(\alpha g_a^u\right)$$

where $\alpha = \lfloor g^{ab} \rfloor_p$ and $g_a = g^a$. Thus we have a special case of the **EHNP**.

Unfortunately the paper [8] has a minor gap in the proof of Theorem 2 of that paper. It is claimed that if g is a primitive root (that is, if $\tau = p-1$) then the obtained problem is exactly the **HNP**. However, this is true only if g_a is a primitive root as well, thus if $\gcd(a, p-1) = 1$.

To fix this gap and to extend the result to the case of $\tau < p-1$, Gonzalez Vasco and Shparlinski [21] have used the bounds of exponential sums from [38] which we have presented in Theorem 10 and Theorem 11.

Using (6.5) we see that under the conditions of Theorem 10 and Theorem 11 the sequence g^x, $x = 0, \dots, T-1$, is $p^{-\delta}$-homogeneously distributed modulo p.

Combining this result with the above arguments and Theorem 19, one can obtain the following statement about the bit security of the Diffie–Hellman secret key.

For each integer $\ell \geq 1$ define the oracle $\mathcal{DH}_\ell$ as an 'black box' which given the values of $X = \lfloor g^x \rfloor_p$ and $Y = \lfloor g^y \rfloor_p$ outputs the value of $\mathrm{MSB}_{\ell,p}(g^{xy})$.

Theorem 20 *Let Q be a sufficiently large integer. The following statement holds with $\vartheta = 1/3$ for all primes $p \in [Q, 2Q]$, and with $\vartheta = 0$ for all primes $p \in [Q, 2Q]$ except at most $Q^{5/6+\varepsilon}$ of them. Let $k = \left\lceil \log^{1/2} p \right\rceil + \lceil \log \log p \rceil$. For any $\varepsilon > 0$, sufficiently large p and any element $g \in \mathbb{F}_p^*$ of multiplicative order $T \geq p^{\vartheta+\varepsilon}$, there exists a probabilistic polynomial time algorithm which for any pair $(a, b) \in [0, T-1]^2$, given the values of $A = \lfloor g^a \rfloor_p$ and $B = \lfloor g^b \rfloor_p$, makes $O\left(\log^{1/2} p\right)$ calls of the oracle $\mathcal{DH}_k$ and computes $\lfloor g^{ab} \rfloor_p$ correctly with probability $1 + O\left(2^{-\log^{1/2} p}\right)$.*

6.2.5 Attack on the Digital Signature Algorithm

On the other hand, in some cases the corresponding exponential sums are new and require a separate study. For example, in [53] the sequence arising in the attack on the Digital Signature Algorithm (DSA) has been studied. We recall the DSA settings. Assume that q and p are primes with $q|p-1$ and that $g \in \mathbb{F}_p$ is a fixed element of multiplicative order q. Let $\mathcal{M}$ be the set of messages to be signed and let $h : \mathcal{M} \to \mathbb{F}_q$ be an arbitrary hash-function. They all (that is, p, q, g, $\mathcal{M}$, h) are *publicly* known.

The *secret key* is an element $\alpha \in \mathbb{F}_q^*$ which is known only to the signer.

To sign a message $\mu \in \mathcal{M}$, the signer chooses a random integer $k \in \mathbb{F}_q^*$ usually called the *nonce*, and which must be kept secret. We define the following two elements of $\mathbb{F}_q$:

$$r(k) = \left\lfloor \lfloor g^k \rfloor_p \right\rfloor_q, \qquad s(k,\mu) = \left\lfloor k^{-1}\left(h(\mu) + \alpha r(k)\right)\right\rfloor_q.$$

The pair $(r(k), s(k,\mu))$ is the *DSA signature* of the message μ with a nonce k.

The attack on the DSA which has been developed in [52] (and which simplifies and improves the attack from [31]) is based on the solving the **HNP** with the sequence

$$t(k,\mu) = \left\lfloor 2^{-\ell} r(k) s(k,\mu)^{-1}\right\rfloor_q, \qquad (k,\mu) \in \mathcal{S}, \tag{6.6}$$

where $\mathcal{S}$ is the set of pairs $(k,\mu) \in [1, q-1] \times \mathcal{M}$ with $s(k,\mu) \neq 0$.

Denote by W the number of solutions of the equation $h(\mu_1) = h(\mu_2)$, $\mu_1, \mu_2 \in \mathcal{M}$. Thus $W/|\mathcal{M}|^2$ is probability of collision and expected to be of order q^{-1} for any practically usable hash function.

In [54] the heuristic results of [52] have been made rigorous.

The central problem is bounding the exponential sums

$$T(c) = \sum_{(k,\mu)\in\mathcal{S}} \mathbf{e}_q\left(ct(k,\mu)\right)$$

where $\mathcal{S}$ the set of pairs $(k,\mu) \in [1, q-1] \times \mathcal{M}$ with $s(k,\mu) \neq 0$ (that is, the set of pairs (k,μ) for which $t(k,\mu)$ is defined).

The following bound of these sums uses

- bounds of exponential sums with exponential functions of S. V. Konyagin and I. E.Shparlinski [38] given by Theorems 10 and 11;
- **Weil's** bound given by Theorem 9;
- **Vinogradov's** method of estimates of double sums [77, 78].

The **main difficulty** is that the double reduction erases any number theoretic structure among the values of $r(k)$.

Theorem 21 *Let Q be a sufficiently large integer. The following statement holds with $\vartheta = 1/3$ for all primes $p \in [Q, 2Q]$, and with $\vartheta = 0$ for all primes $p \in [Q, 2Q]$ except at most $Q^{5/6+\varepsilon}$ of them. For any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $g \in \mathbb{F}_p$ of multiplicative order $q \geq p^{\vartheta+\varepsilon}$ the sequence (6.6) the bound*

$$\max_{\gcd(c,q)=1} |T(c)| = O\left(W^{1/2} q^{3/2-\delta}\right)$$

holds.

Proof. For $\lambda \in \mathbb{F}_q$ we denote by $H(\lambda)$ the number of $\mu \in \mathcal{M}$ with $h(\mu) = \lambda$. We also define the integer $a \in [1, q-1]$ by the congruence $a \equiv 2^{-\ell} c_0 \pmod q$. We have

$$|T(c)| \le \sum_{\lambda \in \mathbb{F}_q} H(\lambda) \left| \sum_{\substack{k=1 \\ \alpha r(k) \not\equiv -\lambda \pmod q}}^{q-1} \mathbf{e}_q\left(a \frac{kr(k)}{\lambda + \alpha r(k)}\right) \right|.$$

Applying the Cauchy inequality we obtain

$$\begin{aligned} |T(c)|^2 \le & \sum_{\lambda \in \mathbb{F}_q} H(\lambda)^2 \\ & \times \sum_{\lambda \in \mathbb{F}_q} \left| \sum_{\substack{k=1 \\ \alpha r(k) \not\equiv -\lambda \pmod q}}^{q-1} \mathbf{e}_q\left(a \frac{kr(k)}{\lambda + \alpha r(k)}\right) \right|^2. \end{aligned}$$

The second sum does not depend on h anymore! (Vinogradov's trick) First of all we remark that

$$\sum_{\lambda \in \mathbb{F}_q} H(\lambda)^2 = W. \tag{6.7}$$

Furthermore,

$$\begin{aligned} & \sum_{\lambda \in \mathbb{F}_q} \left| \sum_{\substack{k=1 \\ \alpha r(k) \not\equiv -\lambda \pmod q}}^{q-1} \mathbf{e}_q\left(a \frac{kr(k)}{\lambda + \alpha r(k)}\right) \right|^2 \\ & \qquad = \sum_{\lambda \in \mathbb{F}_q} \sum_{\substack{k=1 \\ \alpha r(k) \not\equiv -\lambda \pmod q}}^{q-1} \sum_{\substack{m=1 \\ \alpha r(m) \not\equiv -\lambda \pmod q}}^{q-1} \mathbf{e}_q\left(aF_{k,m}(\lambda)\right) \\ & \qquad = \sum_{k,m=1}^{q-1} \sum_{\lambda \in \mathbb{F}_q}{}^{*} \mathbf{e}_q\left(aF_{k,m}(\lambda)\right), \end{aligned}$$

where

$$F_{k,m}(X) = \frac{kr(k)}{X + \alpha r(k)} - \frac{mr(m)}{X + \alpha r(m)}$$

and the symbol $\sum^*$ means that the summation in the inner sum is taken over all $\lambda \in \mathbb{F}_q$ with

$$\lambda \not\equiv -\alpha r(k) \pmod q, \qquad \lambda \not\equiv -\alpha r(m) \pmod q,$$

thus Theorem 9 applies (Weil's bound) **unless** $F_{k,m}(X)$ is constant in $\mathbb{F}_q$.
The function $F_{k,m}(X)$ is constant only if $k = m$ or $r(k) = r(m) = 0$ (in which case we use the trivial bound).
The condition $r(k) = 0$ is equivalent to

$$g^k \equiv qx \pmod p, \qquad k \in [1, q-1],\ x \in [0, L].$$

where

$$L = \left\lfloor \frac{p-1}{q} \right\rfloor.$$

Using Theorems 10 and 11 and the method of proof of Theorem 5, one can now prove that under the conditions of Theorem 21 the last congruence has $O(q^{1-\delta})$ solutions for some $\delta > 0$.
There are also q pairs $k = m$.
Putting everything together gives

$$|T(c)|^2 = O\left(W\left(q^2 \cdot q^{1/2} + (q + q^{2-2\delta})q\right)\right)$$

and the desired result follows. □
Using (6.5) we see that under the conditions of Theorem 21 the sequence (6.6) is $q^{-\delta/3}$-homogeneously distributed modulo q provided that

$$W \le \frac{(\#\mathcal{M})^2}{q^{1-\delta}}. \tag{6.8}$$

This result is based on a combination of the bounds of exponential sums with exponential functions from [38] given in Theorem 10 and Theorem 11, with the *Weil* bound, see [45] and the Vinogradov method of estimates of double sums. As we have mentioned, the inequality (6.8) usually holds in the stronger form $W = O\left(|\mathcal{M}|^2/q\right)$.
Then the above arguments together with Theorem 19 imply the following statement.
For an integer ℓ we define the oracle $\mathcal{DSA}_\ell$ which, for any given DSA signature $(r(k), s(k, \mu))$, $k \in [0, q-1]$, $\mu \in \mathcal{M}$, returns the ℓ least significant bits of k.

Theorem 22 *Let Q be a sufficiently large integer. The following statement holds with $\vartheta = 1/3$ for all primes $p \in [Q, 2Q]$, and with $\vartheta = 0$ for all primes $p \in [Q, 2Q]$ except at most $Q^{5/6+\varepsilon}$ of them. For any $\varepsilon > 0$ there exists $\delta > 0$ such that for any element $g \in \mathbb{F}_p$ of multiplicative order q, where $q \ge p^{\vartheta+\varepsilon}$ is prime, and any hash function h satisfying (6.8), given an oracle*

$\mathcal{DSA}_\ell$ with $\ell = \left\lceil \log^{1/2} q \right\rceil + \lceil \log\log q \rceil$, there exists a probabilistic polynomial time algorithm to recover the DSA secret key α, from $O\left(\log^{1/2} q\right)$ signatures $(r(k), s(k,\mu))$ with $k \in [0, q-1]$ and $\mu \in \mathcal{M}$ selected independently and uniformly at random. The probability of success is at least $1 - 2^{-(\log\log q)\log^{1/2} q}$.

The same result holds for most significant bits and (in a marginally weaker form) for bit strings in the middle.

Practically: Numerical experiments, see [52, 53] show that

- 4 bits of k are always enough,
- 3 bits are often enough,
- 2 bits are possibly enough as well.

Moral

1. Do not use **small** k (to cut the cost of exponentiation in $r(k)$). This can be very tempting because there is no $K^{1/2}$-attack (Shank's *Baby–step–Giant–step* and Pollard's *Rho* do not apply) on $r(k)$ with $k \in [0, K]$.
2. Protect your software/hardware against **timing/power attacks** when the attacker measures the time/power consumption and selects the signatures for which this value is smaller than "on average" – these signatures are likely to correspond to small values of k (because they correspond to faster exponentiation in $r(k)$, timing for other parts of the algorithm is about the same for all k and μ).
3. Use quality **pseudorandom number generators** to generate k, biased generators are dangerous.
4. Do not use **Arazi's cryptosystem** which combines DSA and Diffie-Hellman key exchange protocol – it leakes some bits of k (has been just noticed by *Don Brown & Alfred Menezes*, 2001).
5. Do not buy CryptoLib from **AT&T**: it always uses odd values of k thus one bit is leaked immediately, one more and .... This was observed by Daniel Bleichenbacher and actually was the main motivation for studying this problem.

6.2.6 Other Applications and Open Questions

The method of the proof of Theorem 20 can be used to establish the bit security of several other exponentiation based cryptographic algorithms. Several such schemes, including the *ElGamal cryptosystem* (see Section 8.4 in [50]) and the *Shamir message passing scheme* (see Protocol 12.22 of [50]), have been outlined in [8, 9]. As yet another example we also mention the *Matsumoto–Takachima–Imai key-agreement protocol*, see Section 12.6 of [50]. In fact the treatment of the Shamir message passing scheme in [8] has the same gap as the treatment of the Diffie-Hellman scheme. Accordingly, using exponential sums this gap has been fixed in [21].

In [68] several results on the recently introduced in [42, 43] the *XTR cryptosystem*. However these results are substantially weaker than those known for the aforementioned. The main reason for this is that in studying the XTR the corresponding character sums are over small subgroups of *extension* fields arise and for such sums there is no analogue of Theorem 10 and Theorem 11. Accordingly, the paper [68] uses a different way of estimating the distribution of multipliers t of the corresponding **EHNP**. Unfortunately this leads to a substantially weaker result. To be more precise, to apply an analogue of the approach of Section 6.2.4 the XTR one needs to improve the bounds of the exponential sums

$$\max_{\gamma\in\mathbb{F}_{p^6}^*}\left|\sum_{t\in\mathcal{G}}\exp\left(2\pi i \mathrm{Tr}\left(\gamma t\right)/p\right)\right| \le p^3$$

where $\mathrm{Tr}(z) = z + z^p + \ldots + z^{p^5}$ is the trace of $z \in \mathbb{F}_{p^6}$ in $\mathbb{F}_p$ and $\mathcal{G}$ is a subgroup of $\mathbb{F}_{p^6}^*$, see Theorem 8.78 in [45] (combined with Theorem 8.24 of the same work) or the bound (3.15) in [38]. This bound is trivial for $\#\mathcal{G} \le p^3$ while the subgroups relevant to XTR are of size of order p^2. Thus in [68] an alternative approach has been used which is based on the fact, even if the sequence $\mathcal{T}$ is not known to be homogeneously distributed but at least admits a non-trivial upper bound for the number of its elements in an interval one can still obtain some analogues of (6.4). Then the upper bound from [13] on the number of zeros of sparse polynomials can used to extract such information.

The result of Theorem 22 has been extended to other DSA-like signature schemes, including the *elliptic curve* version of DSA in [16, 54]. In particular, the bound of [37] provides an analogue of Theorem 10 for exponential sums over an orbit generated by a point on an elliptic curve, see [54]. However some interesting questions still remain open. For example, for the *Nyberg–Rueppel* signature scheme the range of p and q in which the results of [16] are nontrivial are narrower than in practical applications. It is shown in [16] that the attack designed in that paper on the Nyberg–Rueppel signature scheme can be reduced to **EHNP** with the sequence of multipliers

$$r(k,\mu) = \left\lfloor \left\lfloor h(\mu) g^k \right\rfloor_p \right\rfloor_q, \qquad (k,\mu) \in [1, q-1] \times \mathcal{M}.$$

Unfortunately it is not clear how to estimate the exponential sums

$$\sum_{\mu \in \mathcal{M}} \sum_{k \in \mathbb{F}_q^*} \exp\left(2\pi i c r(k,\mu)\right), \qquad c \in [1, q-1],$$

and obtaining such a bound is an interesting open question. Using a rather indirect method, it has been shown in [16] that the sequence $r(k,\mu)$ is $2^{-\log^{1/2} q}$ homogeneously distributed modulo q, provided that

$$W \le \frac{(\#\mathcal{M})^2 q^{3-\delta}}{p^3}$$

for some $\delta > 0$. We remark that in the settings of the Nyberg–Rueppel signature scheme it is natural to assume that h is bijective, that is, $W = \#\mathcal{M}$. Also, if the message set $\mathcal{M}$ is "dense" (that is, $\#\mathcal{M}$ is of order p) then the above result holds for $q \ge p^{2/3+\delta}$. It would be very interesting to lower this bound.

Yet another modifications of the **HNP** has recently been introduced in [30]. Namely, that paper introduces the following

> Hidden Number Problem with Hidden Multiplier, **HNP-HM**: *Recover a number $\alpha \in \mathbb{F}_p$ such that for many unknown random $t \in \mathcal{T}$ we are given $\mathrm{MSB}_{\ell,p}(\alpha t)$, $\mathrm{MSB}_{\ell,p}(t)$ and $\mathrm{MSB}_{\ell,p}(\alpha)$ for some $\ell > 0$.*

In the case $\mathcal{T} = \mathbb{F}_p^*$ and $\ell \ge (4/5 + \varepsilon)\log p$ the paper [30] provides a polynomial time algorithm for the **HNP-HM**. In fact it also works in more general residue rings (which is important for applications to [62]). As one can see this result is substantially weaker than those known for **HNP** and **EHNP** where one can take ℓ of order $\log^{1/2} p$. However, using exponential sums, it has been shown in [30] that indeed for **HNP-HM** to have a unique solution the value of ℓ must be very large. Namely for $\ell \le (1/2 + \varepsilon)\log p$ there can be exponentially many possibilities for α.

The aforementioned algorithm has been used in [30] to establish a certain bit security result for the "timed-release crypto" introduced by Rivest, Shamir and Wagner [62] and also to design a "correcting" algorithm for noisy exponentiation black-boxes.

It is an interesting and challenging problem to study **HNP-HM** for more general sequences $\mathcal{T}$, in particular for subgroups of $\mathbb{F}_p^*$.

In the case $\mathcal{T} = \mathbb{F}_p^*$ the paper [9] provides a *non-uniform* polynomial time algorithm for the **HNP** which works with $\ell = O(\log\log p)$. We recall that non-uniformity means that the algorithm exists but to actually design this algorithm one may need exponential time (thus such algorithms are of rather limited value). Nevertheless it would be of interest to extend this result to subgroups of $\mathbb{F}_p^*$. In order to get such a generalisation one needs an analogue of Lemma 2.4 for subgroups and this seems to be a rather feasible task taking into account the bounds of exponential sums of Theorem 10 and Theorem 11.

Finally, several more modifications of the **HNP** have been considered in the papers [7, 20, 36, 71, 76]. However they are of more algebraic than geometric nature and lattices have not been involved in their study.

Bibliography

[1] M. Ajtai, R. Kumar and D. Sivakumar, 'A sieve algorithm for the shortest lattice vector problem' *Proc. 33rd ACM Symp. on Theory of Comput.*, Crete, Greece, July 6-8, 2001, 601–610.

[2] L. Babai, 'On Lovász' lattice reduction and the nearest lattice point problem', *Combinatorica*, **6** (1986), 1–13.

[3] A. M. Barg, 'Incomplete sums, DC-constrained codes, and codes that maintain synchronization', *Designs, Codes and Cryptography*, **3** (1993), 105–116.

[4] A. M. Barg, 'A large family of sequences with low periodic correlation', *Discr. Math.*, **176** (1997), 21–27.

[5] A. M. Barg and S. N. Litsyn, 'On small families of sequences with low periodic correlation', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **781** (1994), 154–158.

[6] L. A. Bassalygo and V. A. Zinoviev, 'Polynomials of special form over a finite field with maximum modulus of the trigonometric sum', *Uspechi Matem. Nauk*, **52** (1997) 2, 31–44 (in Russian).

[7] D. Boneh and I. E. Shparlinski, 'On the unpredictability of bits of the elliptic curve Diffie–Hellman scheme', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2139** (2001), 201–212.

[8] D. Boneh and R. Venkatesan, 'Hardness of computing the most significant bits of secret keys in Diffie–Hellman and related schemes', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1109** (1996), 129–142.

[9] D. Boneh and R. Venkatesan, 'Rounding in lattices and its cryptographic applications', *Proc. 8th Annual ACM-SIAM Symp. on Discr. Algorithms*, ACM, NY, 1997, 675–681.

[10] V. Boyko, M. Peinado and R. Venkatesan, 'Speeding up discrete log and factoring based schemes via precomputations', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1403** (1998), 221–234.

[11] E. Brickell, D.M. Gordon, K.S. McCurley, and D. Wilson, 'Fast exponentiation with precomputation', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **658** (1993), 200–207.

[12] D. A. Burgess, 'The distribution of quadratic residues and non-residues', *Mathematika*, **4** (1957), 106–112.

[13] R. Canetti, J. B. Friedlander, S. V. Konyagin, M. Larsen, D. Lieman and I. E. Shparlinski, 'On the statistical properties of Diffie–Hellman distributions', *Israel J. Math.*, **120** (2000), 23–46.

[14] R. Canetti, J. B. Friedlander and I. E. Shparlinski, 'On certain exponential sums and the distribution of Diffie–Hellman triples', *J. London Math. Soc.*, **59** (1999), 799–812.

[15] P. Deligne, 'La conjecture de Weil, I', *Inst. Hautes Etudes Sci. Publ. Math.*, **43** (1974), 273–307.

[16] E. El Mahassni, P. Q. Nguyen and I. E. Shparlinski, 'The insecurity of Nyberg–Rueppel and other DSA-like signature schemes with partially known nonces', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2146** (2001), (to appear).

[17] J. B. Friedlander, M. Larsen, D. Lieman and I. E. Shparlinski, 'On correlation of binary M-sequences', *Designs, Codes and Cryptography*, **16** (1999), 249–256.

[18] J. B. Friedlander and I. E. Shparlinski, 'On the distribution of Diffie–Hellman triples with sparse exponents', *SIAM J. Discr. Math.*, **14** (2001), 162–169.

[19] C. F. Gauss, *Disquisitiones arithmeticae*, Fleischer, Leipzig, 1801.

[20] M. I. González Vasco, M. Näslund and I. E. Shparlinski, 'The hidden number problem in extension fields and its applications', *Preprint*, 2001, 1–12.

[21] M. I. González Vasco and I. E. Shparlinski, 'On the security of Diffie–Hellman bits', *Proc. Workshop on Cryptography and Computational Number Theory*, Singapore 1999, Birkhäuser, 2001, 257–268.

[22] M. I. González Vasco and I. E. Shparlinski, 'Security of the most significant bits of the Shamir message passing scheme', *Math. Comp.* (to appear).

[23] T. W. Cusick and H. Dobbertin, 'Some new three-valued correlation functions for binary sequences', *IEEE Trans. Inform. Theory*, **42** (1996), 1238–1240.

[24] G.H. Hardy and J. E. Littlewood, 'Some problems of "Partitio Numerorum". I A new solution of Waring's problem', *Göttingen Nachrichten*, 1920, 231–267.

[25] D. R. Heath-Brown and S. Konyagin, 'New bounds for Gauss Sums derived from kth powers, and for Heilbronn's exponential sum', *Ouart. J. Math.*, **51** (2000), 221–235.

[26] T. Helleseth, 'Some results about the cross-correlation function between two maximal linear sequences', *Discr. Math.*, **16** (1976), 209–232.

[27] T. Helleseth,'A note on the cross-correlation function between two binary maximal length linear sequences', *Discr. Math.*, **23** (1978), 301–307.

[28] T. Helleseth, 'On the crosscorrelation of m-sequences and related sequences with ideal autocorrelation', *Proc. Intern. Conf. on Sequences and their Applications (SETA'01)*, Bergen, 2001, Springer-Verlag, (to appear).

[29] T. Helleseth and K. Yang, 'On binary sequences of period $p^m - 1$ with optimal autocorrelation', *Proc. Intern. Conf. on Sequences and their Applications (SETA'01)*, Bergen, 2001, Springer-Verlag, (to appear).

[30] N. A. Howgrave-Graham, P. Q. Nguyen and I. E. Shparlinski, 'Hidden number problem with hidden multipliers, timed-release crypto and noisy exponentiation', *Preprint*, 2001, 1–11.

[31] N. A. Howgrave-Graham and N. P. Smart, 'Lattice attacks on digital signature schemes', *Designs, Codes and Cryptography*, **23** (2001), 283–290.

[32] L.-K. Hua, 'On an exponential sum', *J. Chinese Math. Soc.*, **2** (1940), 301–312.

[33] L.-K. Hua, *Abschätzungen von Exponentialsummen und ihre Anwendung in der Zahlentheorie*, Leipzig, Teubner-Verlag, 1959.

[34] R. Kannan, 'Algorithmic geometry of numbers', *Annual Review of Comp. Sci.*, **2** (1987), 231–267.

[35] R. Kannan, 'Minkowski's convex body theorem and integer programming', *Math. of Oper. Research*, **12** (1987), 231–267.

[36] E. Kiltz, 'A primitive for proving the security of every bit and about universal hash functions & hard core bits', *Preprint*, 2001, 1–19.

[37] D. R. Kohel and I. E. Shparlinski, 'Exponential sums and group generators for elliptic curves over finite fields', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1838** (2000), 395–404.

[38] S. V. Konyagin and I. Shparlinski, *Character sums with exponential functions and their applications*, Cambridge Univ. Press, Cambridge, 1999.

[39] N. M. Korobov, 'On the distribution of digits in periodic fractions', *Math. USSR – Sbornik*, **89** (1972), 654–670 (in Russian).

[40] N. M. Korobov, *Exponential sums and their applications*, Kluwer Acad. Publ., Dordrecht, 1992.

[41] A. K. Lenstra, H. W. Lenstra and L. Lovász, 'Factoring polynomials with rational coefficients', *Mathematische Annalen*, **261** (1982), 515–534.

[42] A. K. Lenstra and E. R. Verheul, 'The XTR public key system', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1880** (2000), 1–19.

[43] A. K. Lenstra and E. R. Verheul, 'Key improvements to XTR', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1976** (2000), 220–233.

[44] V. I. Levenshtein, 'Bounds for packing in metric spaces and certain applications', *Problemy Kibernetiki*, **40** (1983), 44–110 (in Russian).

[45] R. Lidl and H. Niederreiter, *Finite fields*, Cambridge University Press, Cambridge, 1997.

[46] J. E. Littlewood, 'Research in the theory of Riemann ζ-function', *Proc. Lond. Math. Soc.*, **20** (1922) (2), XXII–XXVIII.

[47] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*, North-Holland, Amsterdam, 1977.

[48] Mazur L., 'On some codes correcting asymmetrical errors', *Problemy Peredachi Inform.*, **10** (1974), 40–46 (in Russian).

[49] D. Micciancio, 'On the hardness of the shortest vector problem', *PhD Thesis*, MIT, 1998.

[50] A. J. Menezes, P. C. van Oorschot and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Boca Raton, FL, 1996.

[51] L. J. Mordell, 'On a sum analogous to a Gauss sum', *Quart. J. Math. Oxford*, **3** (1932), 161–167.

[52] P. Q. Nguyen, 'The dark side of the hidden number problem: Lattice attacks on DSA', *Proc. Workshop on Cryptography and Computational Number Theory*, Singapore 1999, Birkhäuser, 2001, 321–330.

[53] P. Q. Nguyen and I. E. Shparlinski, 'The insecurity of the Digital Signature Algorithm with partially known nonces', *J. Cryptology* (to appear).

[54] P. Q. Nguyen and I. E. Shparlinski, 'The insecurity of the elliptic curve Digital Signature Algorithm with partially known nonces', *Preprint*, 2000, 1–24.

[55] P. Q. Nguyen, I. E. Shparlinski and J. Stern, 'Distribution of modular sums and the security of the server aided exponentiation', *Proc. Workshop on Cryptography and Computational Number Theory*, Singapore 1999, Birkhäuser, 2001, 331–342.

[56] P. Q. Nguyen and J. Stern, 'Lattice reduction in cryptology: An update', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1838** (2000), 85–112.

[57] P. Q. Nguyen and J. Stern, 'The two faces of lattices in cryptology', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2146** (2001), (to appear).

[58] H. Niederreiter, 'Quasi-Monte Carlo methods and pseudo-random numbers', *Bull. Amer. Math. Soc.*, **84** (1978), 957–1041.

[59] H. Niederreiter, *Random number generation and quasi–Monte Carlo methods*, SIAM, Philadelphia, 1992.

[60] H. Niederreiter and I. E. Shparlinski, 'Recent advances in the theory of nonlinear pseudorandom number generators', *Preprint*, 2001, 1–16.

[61] F. Özbudak, 'On lower bounds on incomplete character sums over finite fields', *Finite Fields and Their Appl.*, **2** (1996) 173–191.

[62] R. L. Rivest, A. Shamir and D. A. Wagner, 'Time-lock puzzles and timed-release crypto', *Preprint*, 1996, 1–9.

[63] F. Rodier, 'Minoration de certain sommes exponentielles, 2', *Arithmetic, Geometry and Coding Theory*, Walter de Gruyter, Berlin, 1996, 185–198.

[64] C. P. Schnorr, 'A hierarchy of polynomial time basis reduction algorithms', *Theor. Comp. Sci.*, **53** (1987), 201–224.

[65] I. E. Shparlinski, *Finite fields: Theory and computation*, Kluwer Acad. Publ., Dordrecht, 1999.

[66] I. E. Shparlinski, *Number theoretic methods in cryptography: Complexity lower bounds*, Birkhauser, Basel, 1999.

[67] I. E. Shparlinski, 'Sparse polynomial approximation in finite fields', *Proc. 33rd ACM Symp. on Theory of Comput.*, Crete, Greece, July 6-8, 2001, 209–215.

[68] I. E. Shparlinski, 'On the generalised hidden number problem and bit security of XTR', *Proc. the 14th Symp. on Appl. Algebra, Algebraic Algorithms, and Error-Correcting Codes*, Melbourne, 2001, Springer-Verlag, Berlin (to appear).

[69] I. E. Shparlinski, 'On the uniformity of distribution of the RSA pairs', *Math. Comp.*, **70** (2001), 801–808.

[70] I. E. Shparlinski, 'Security of most significant bits of g^{x^2}', *Preprint*, 2001, 1–7.

[71] I. E. Shparlinski, 'Security of polynomial transformations of the Diffie–Hellman key', *Preprint*, 2001, 1–8.

[72] S. B. Stečkin, 'An estimate of a complete rational exponential sum', *Proc. Math. Inst. Acad. Sci. USSR*, Moscow, **143** (1977), 188–207 (in Russian).

[73] S. A. Stepanov, 'Character sums and coding theory', *Finite Fields and Applications*, London Math. Soc. Lect., Notes Ser., Vol. 233, Cambridge Univ. Press, Cambridge, 1996, 355–378.

[74] S. A. Stepanov, 'Character sums, algebraic curves and coding theory', *Lect. Notes in Pure and Appl. Math.*, Marcel Dekker, NY, **193** (1997), 313–345.

[75] R. C. Vaughan, *The Hardy–Littlewood method*, Cambridge Univ. Press, Cambridge, 1981.

[76] E. R. Verheul, 'Certificates of recoverability with scalable recovery agent security', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1751** (2000), 258–275.

[77] I. M. Vinogradov, 'On Weyl's sums', *Matem. Sbornik*, **42** (1935), 258–275 (in Russian).

[78] I. M. Vinogradov, 'Representation of an odd number as a sum of three primes', *Doklady Russian Acad. Sci.*, **15** (1937), 291–294 (in Russian).

[79] A. Weil, 'On some exponential sums', *Proc. Nat. Sci. Acad. Sci U.S.A.*, **34** (1948), 204-207.

[80] H. Weyl, 'Über die Gleichverteilung von Zahlen mod Eins', *Math. Ann.*, **77** (1916), 313–352.