# The Prime Number Theorem and Landau's Extremal Problems

Szilárd Gy. Révész

## Contents

## §0. Introduction

A century ago Edmund Landau analyzed the then recent proofs of Hadamard and, in particular, of de la Vallée Poussin for the Prime Number Theorem. As de la Vallée Poussin proved also an error term

$$\pi(x) = \mathrm{li}(x) + O(xe^{-c\sqrt{x}}), \tag{1}$$

Landau's idea was to improve upon the error term optimizing the technical tools employed. He presented in his "Handbuch..." two attempts. Both methods are technically by now standard, but still lengthy and full of calculations. The form of error terms deducible present themselves as

$$\pi(x) = \int_2^x \frac{dt}{\log t} + O\big(x \cdot \exp(-K \log^L x)\big), \tag{2}$$

and Landau sought the optimal values of $L$ an $K$ deducible using the method of de la Vallée Poussin; more precisely, optimizing use of the positive trigonometric polynomials occupying a key role in the original proof.

## §1. A few properties of the Riemann zeta function

In this section we deal with the Riemann zeta function

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \left( \frac{1}{1-\frac{1}{p^s}} \right) \qquad (\sigma := \Re s > 1). \tag{3}$$

Here the first, defining sum – the *Dirichlet series* of the identically 1 sequence – is absolutely and locally uniformly convergent in $\sigma > 1$. Thus in the whole domain of convergence $\sigma > 1$, $\zeta(s)$ is analytic, each function $n^{-s} = \exp(-s \log n)$ being an entire function in $s$. Moreover, by absolute convergence and the unique prime factorization, the second form – the *Euler product* form – follows everywhere in $\sigma > 1$. In fact, this Euler product form is often called the *analytic form of the unique prime factorization theorem*, for obvious reasons. The same way as the sum, it also gives a locally uniformly convergent representation: $\zeta(s)$ is the limit of partial products of analytic functions $(1-p^{-s})^{-1}$. However, here the analytic functions used are not entire functions, but only regular for $\sigma > 0$; and neither representation is convergent, hence is not valid, for any larger halfplane than $\sigma > 1$.

Still, the Riemann zeta function $\zeta(s)$ itself can be meromorphically continued all over the complex plane $\mathbb{C}$. Also, the Rieman zeta function admits a functional equation

$$\zeta(s) = 2^s \pi^{s-1} \Gamma(1-s) \sin(s\pi/2) \zeta(1-s),$$

or, in a more symmetric form,

$$\xi(s) = \xi(1-s) \qquad \text{with} \quad \xi(s) := (s-1)\Gamma(s/2+1)\pi^{-s/2}\zeta(s). \tag{4}$$

But what is the definition of $\zeta(s)$ for $\sigma \leq 1$ in these formulae? The functional equation makes sense if only we have some extended definition for $\zeta(s)$ for values out of the halfplane of convergence of (3). This can be accomplished by either defining the analytic continuation along the way of proving the functional equation, or by extending the function first and treating the functional equation only after it.

First we present a direct analytic continuation by means of partial summation, as this is most suitable to derive the basic estimates we need along the course of our work. By partial summation the formula

$$\sum_{a<n\leq b} f(n)g(n) = F(b)g(b) - F(a)g(a) - \int_a^b F(t)g'(t)dt \tag{5}$$

$$\left( F(x) := \sum_{n<x} f(n), \quad g \in C^1[a,b] \right)$$

is meant; in fact this is another formulation of partial integration for Stieltjes integrals with respect to $dF$. As a technical tool, we denote

$$r(x) := \int_1^x \left(\{y\} - \frac{1}{2}\right) dy \in \left[\frac{-1}{8}, 0\right] \quad (\forall x \in \mathbb{R}), \qquad r(N) = 0 \quad (\forall N \in \mathbb{N}). \tag{6}$$

**Proposition 1.** *The Riemann zeta function continues analytically to $\sigma > -1$, except for one singularity at $s = 1$ where it has a first order pole of residuum 1. Furthermore, $\zeta(s)$*

i) *has the form*

$$\zeta(s) = \frac{1}{s-1} + \frac{1}{2} - s(s+1) \int_1^\infty \frac{r(x)dx}{x^{2+s}},$$

*with $r(x)$ defined in (6),*

ii) *satisfies*

$$\zeta(s) = \sum_{n=1}^N \frac{1}{n^s} - \frac{1}{2}N^{-s} + \frac{1}{s-1}N^{1-s} - s(s+1) \int_N^\infty \frac{r(x)}{x^{2+s}}dx, \tag{7}$$

*for all $N \in \mathbb{N}$ and for $\sigma > -1$*

iii) *satisfies for $\sigma > -1$*

$$\left|\zeta(s) - \frac{s+1}{2(s-1)}\right| \le \frac{|s(s+1)|}{8(\sigma+1)},$$

iv) *and satisfies also $\frac{1}{\sigma-1} < \zeta(\sigma) < \frac{\sigma}{\sigma-1}$ for all $1 \le \sigma$.*

*Proof.* Recall that $\zeta(s) = \lim_{N\to\infty} \sum_{n=1}^N \frac{1}{n^s}$ in $\sigma > 1$. Putting $a = 1/2$ and $b = N$, say, partial summation (5) leads to

$$\begin{aligned}
\sum_{n=1}^N \frac{1}{n^s} &= \left(\sum_{n=1}^N 1\right) N^{-s} - \int_1^N [x] \frac{-s}{x^{s+1}} dx = N^{1-s} + s \int_1^N \frac{(x - \{x\})}{x^{s+1}} dx \\
&= N^{1-s} + s \int_1^N \frac{1}{x^s} dx + \frac{1}{2} \int_1^N \frac{-s}{x^{s+1}} dx - s \int_1^N \frac{\left(\{x\} - \frac{1}{2}\right)}{x^{s+1}} dx \\
&= N^{1-s} + \left[\frac{sx^{1-s}}{1-s}\right]_1^N + \left[\frac{x^{-s}}{2}\right]_1^N - s \int_1^N \frac{\left(\{x\} - \frac{1}{2}\right)}{x^{s+1}} dx \\
&= \frac{s}{s-1} - \frac{1}{2} + \frac{1}{2}N^{-s} - \frac{1}{s-1}N^{1-s} - s \int_1^N \frac{\left(\{x\} - \frac{1}{2}\right)}{x^{s+1}} dx.
\end{aligned} \tag{8}$$

Observe that the last integral in (8) converges absolutely and locally uniformly for $\sigma > 0$. Taking limits in $N \to \infty$, we get the formula

$$\zeta(s) = \frac{s}{s-1} - \frac{1}{2} - s\int_1^\infty \frac{\left(\{x\} - \frac{1}{2}\right)}{x^{s+1}} dx, \tag{9}$$

providing an analytic continuation of $\zeta(s)$ all over the right halfplane $\Re s > 0$, as the integral is locally uniformly and absolutely convergent there. Let us now apply repeated partial integration in (8) to get

$$\begin{aligned}\sum_{n=1}^N \frac{1}{n^s} &= \frac{s}{s-1} - \frac{1}{2} + \frac{1}{2}N^{-s} - \frac{1}{s-1}N^{1-s} - s\left\{\left[\frac{r(x)}{x^{1+s}}\right]_1^N - \int_1^N r(x)\frac{(-1-s)}{x^{2+s}}dx\right\} \\ &= \frac{s}{s-1} - \frac{1}{2} + \frac{1}{2}N^{-s} - \frac{1}{s-1}N^{1-s} - s\left\{0 + \int_1^N r(x)\frac{1+s}{x^{2+s}}dx\right\},\end{aligned} \tag{10}$$

where $r(x)$ is defined in (6). The last integral in (10) converges absolutely and locally uniformly even for $\sigma > -1$. Letting $N \to \infty$, we get the formula

$$\zeta(s) = \frac{s}{s-1} - \frac{1}{2} - s(s+1)\int_1^\infty \frac{r(x)dx}{x^{2+s}} = \frac{1}{s-1} + \frac{1}{2} - s(s+1)\int_1^\infty \frac{r(x)dx}{x^{2+s}}, \tag{11}$$

first following from (10) only for $\sigma > 1$. However, (11) extends even to $\sigma > -1$ analytically. This defines the analytic (more precisely, meromorphic) continuation of the Riemann zeta function, and shows its meromorphic behavior as having one, simple pole at $s = 1$ and behaving analytically everywhere else in $\sigma > -1$.

Comparing (10) and (11) it is immediate that we also have (7), which, in fact, contains only terms analytic in $\sigma > -1$, except for the pole at $s = 1$ of $N^{1-s}/(s-1)$. Observe that (11) implies also part (iii) of the assertion if we take into account

$$\left|\int_1^\infty \frac{r(x)dx}{x^{2+s}}\right| \leq \frac{1}{8}\int_1^\infty \frac{dx}{x^{2+\sigma}} = \frac{1}{8(1+\sigma)}.$$

Moreover, for real $1 < s = \sigma$ comparison of the integral $\int_1^\infty x^{-\sigma}dx$ with the defining sum of $\zeta(s)$ yields part (iv) of the statement. □

Already Riemann has calculated the first few zeros of $\zeta(s)$, and found that the first complex zero occurs at $1/2 + i14, \ldots$. That is, $\zeta(s)$ does not vanish in the close vicinity of the real line (in $\sigma \geq 0$). For us it suffices to see the following.

**Proposition 2.** *The Riemann zeta function $\zeta(s)$ does not vanish in $1/2 \leq \sigma \leq 1$, $-\sqrt{5} \leq t \leq \sqrt{5}$.*

*Proof.* We simply use the elementary formula obtained as part (iii) of the previous Proposition 1. If $\zeta(s)$ vanishes, then this gives $4(\sigma+1) \leq |s(s-1)|$. Let us fix $t$ and assume, as we may, that $0 \leq t$. On the line segment $[1/2 + it, 1 + it]$ we have $|s(s-1)| \leq |s|^2 = t^2 + \sigma^2 \leq t^2 + \sigma$, hence in the full range $1/2 \leq \sigma \leq 1$ we must certainly have $4(\sigma+1) \leq t^2 + \sigma$, that is, $t^2 \geq 4 + 3\sigma \geq 11/2 > 5$. □

**Proposition 3.** *The Riemann zeta function also satisfies*

i) $|\zeta(s)| \leq c \log |t| \qquad |t| \geq 2,\ 1 - \frac{1}{\log |t|} \leq \sigma$ *and*

ii) $|\zeta'(s)| \leq c \log^2 |t| \qquad |t| \geq 2,\ 1 - \frac{1}{\log |t|} \leq \sigma.$

iii) $|\zeta(s)| \leq c|t|^{1-\sigma}/(1-\sigma)$ *for all* $|t| \geq 2$, $-1 \leq \sigma < 1$.

iv) $|\zeta(s)| \leq c|t|^{(1-\sigma)_+} \log |t|$ *for all* $|t| \geq 2$, $-1 \leq \sigma$.

*Proof.* For $\sigma \geq 2$ we have uniform boundedness of $\zeta(s)$ and $\zeta'(s)$, in view of absolute and uniform majorization of their series representation by $\zeta(2)$ and $\zeta'(2)$, respectively. So assume $\sigma < 2$. We also assume, as we may, $t > 0$, hence $t \geq 2$. On applying (7) with $N := [t] + 1$, say, we obtain

$$
\begin{aligned}
|\zeta(s)| \leq & \max\{1, N^{-\sigma}\} + \int_1^N x^{-\sigma} dx + N^{-\sigma} + \frac{N^{-(\sigma-1)}}{t} + (t+2)(t+3)\frac{N^{-1-\sigma}}{1+\sigma} \\
\leq & 1 + cN^{1-\sigma} + \frac{N^{1-\sigma} - 1}{1-\sigma} + c\left(\frac{t}{N}\right)^2 N^{1-\sigma} \ll \frac{t^{1-\sigma} - 1}{1-\sigma} + 1 + t^{1-\sigma}. \qquad (12)
\end{aligned}
$$

Observe that we already have (iii). Moreover, we have for $\sigma < 1$

$$
\frac{t^{1-\sigma} - 1}{1-\sigma} \leq \max_{0 \leq \theta \leq 1-\sigma} \frac{d}{d\theta} e^{\theta \log t} = \log t\ t^{1-\sigma} \qquad (13)
$$

and for $\sigma > 1$

$$
\frac{t^{1-\sigma} - 1}{1-\sigma} = \frac{1 - t^{-(\sigma-1)}}{\sigma - 1} \leq \max_{0 \leq \theta \leq \sigma-1} \frac{d}{d\theta} e^{-\theta \log t} = \log t, \qquad (14)
$$

which proves (iv) for $\sigma \neq 1$ and whence by continuity for all $\sigma > -1$, too. Consider now the domain $\sigma \geq 1 - 2/\log t$ and $t \geq 3/2$. In this domain $t^{1-\sigma} \leq t^{2/\log t} = e^2$, hence also (i) obtains. In fact, we got more (i.e., a larger domain) than we need for part (i), but it makes also the next step easy. Indeed, for an application of Cauchy's coefficient estimate at $s = \sigma + it$ in a circle of radius $\rho := 1/4 \log t$, the circle stays in this domain when $t \geq 2$ and $\sigma \geq 1 - 1/\log t$. That proves the second part, too. Alternatively, the similar partial summation formula (5) can be applied even to $\zeta'(s)$; or, formula (7) can be differentiated and then estimated. $\square$

In our work it is important to observe that some estimates, valid to the right of the line $\sigma = 1$, extend on and to the left of it in a certain range.

**Proposition 4.** *Assume that with certain constants $c > 0$, and $b \geq 1$, and with a constant $a > 0$ small enough, the Riemann zeta function satisfies for all $|t| \geq 2$ the estimate*

$$
|\zeta(s_0)| \geq c \log^{-b} |t| \quad \text{for} \quad s_0 = \sigma_0 + it \quad \text{with} \quad \sigma_0 := \sigma_0(t) := 1 + \frac{a}{\log^{b+2} |t|}.
$$

*Then we must have*

$$|\zeta(s)| \geq \frac{c/2}{\log^b |t|} \quad \text{for all} \quad s = \sigma + it \quad \text{with} \quad 2 > \sigma > 1 - \frac{a}{\log^{b+2} |t|}.$$

*Furthermore, we must also have*

$$\left| \frac{\zeta'(s)}{\zeta(s)} \right| \ll \log^{b+2} |t| \quad \text{for all} \quad s = \sigma + it \quad \text{with} \quad 2 > \sigma > 1 - \frac{a}{\log^{b+2} |t|}. \tag{15}$$

*Proof.* Again, assume $t \geq 2$. Since $\zeta(s)$ is analytic, clearly we can write

$$|\zeta(s)| \geq |\zeta(s_0)| - |\zeta(s_0) - \zeta(s)| \geq c \log^{-b} t - \max_{\sigma} |\zeta'(\sigma + it)| \frac{2a}{\log^{b+2} t}.$$

Since $\zeta'$ can be estimated according to part (ii) of Proposition 2, we now get with $C$ denoting the constant there

$$|\zeta(s)| \geq c \log^{-b} t - C \log^2 t \frac{2a}{\log^{b+2} t} > \frac{c/2}{\log^b t}.$$

if $a < c/(4C)$. The first part of the assertion follows, while using the estimate of Proposition 3 (ii) for $\zeta'$, even the second part obtains. $\square$

## §2. The logarithmic derivative of $\zeta(s)$ in the critical strip

We have seen that $\zeta(s)$ has an analytic (meromorphic) continuation. Even if $\zeta(s)$ may vanish at some discrete points, $1/\zeta(s)$ is meromorphic together with $\zeta(s)$ at least in the same halfplane $\sigma > -1$. Moreover, it is analytic for $\sigma > 1$, as $\zeta(s) \neq 0$ in $\sigma > 1$. This follows from the locally uniformly convergent Euler product representation (3), as from that we get

$$\frac{1}{\zeta(s)} = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s},$$

where $\mu(n) = (-1)^k$ if $n$ is squarefree with $k$ distinct prime divisors, and $\mu(n) = 0$ if $n$ is not squarefree. (This arithmetical function is Möbius's function.) It is immediate that

$$\left|\frac{1}{\zeta(s)}\right| \leq \sum_{n=1}^{\infty} \frac{1}{n^\sigma} = \zeta(\sigma) \qquad (\sigma > 1), \tag{16}$$

an even sharper form of nonvanishing of $\zeta(s)$ in $\sigma > 1$. However, to estimate order of magnitude of $1/\zeta(s)$ is quite intricate. For our needs here we analyze the logarithmic derivative $\zeta'/\zeta$, which is defined meromorphically as the quotient of $\zeta'$ and $\zeta$ (even if defining also the logarithm would take more effort and complications).

Let us recall first a fundamental formula of Jensen.

**Lemma 1 (Jensen).** *Let $f(z)$ be analytic throughout a disk $|z - z_0| \leq R$ and let the zeroes of $f$ in the disk be $z_1, \dots, z_n$, all listed according to multiplicity. Assume that $f(z_0) \neq 0$, i.e. $z_0 \neq z_j$ for $j = 1, \dots, n$. Then we have*

$$\log\left|f(z_0)\frac{R}{z_1}\cdots\frac{R}{z_n}\right| = \frac{1}{2\pi}\int_0^{2\pi} \log\left|f(Re^{i\varphi} + z_0)\right| d\varphi. \tag{17}$$

*Proof.* If $f \neq 0$ throughout the disk, i.e. $n = 0$, then even $\log f$ is analytic and Cauchy1s formula applies to $\log f$. Since $\Re \log w = \log|w|$, taking the real part of Cauchy's formula gives

$$\log|f(z_0)| = \frac{1}{2\pi}\int_0^{2\pi} \log\left|f(Re^{i\varphi} + z_0)\right| d\varphi, \tag{18}$$

that is, (17) in case $n = 0$. Consider now the case when there are $n \geq 1$ zeroes, and put

$$g(z) := \frac{f(z)}{B(z)}, \quad \text{with} \quad B(z) := \prod_{j=1}^{n} \left(\frac{R(z - z_j)}{R^2 - \overline{(z_j - z_0)}(z - z_0)}\right).$$

The zeroes of $f$ and the Blaschke product $B$ cancel, hence $g \neq 0$ is analytic throughout the disk. An application of (18) for $g$ gives

$$\log|g(z_0)| = \frac{1}{2\pi}\int_0^{2\pi} \log\left|g(Re^{i\varphi}+z_0)\right| d\varphi = \frac{1}{2\pi}\int_0^{2\pi} \log\left|f(Re^{i\varphi}+z_0)\right| d\varphi,$$

since all Blaschke factors map the bounding circle $|z-z_0| = R$ onto $|w| = 1$ and thus $|B(z)|$ does not change the integral on the right. On the other hand it is easy to substitute $z_0$ into $g$ to get

$$\log|g(z_0)| = \log\left|f(z_0)\frac{R}{z_1}\cdots\frac{R}{z_n}\right|,$$

and combining the last two formulas yields (17). □

An obvious corollary is the following estimate of Jensen.

**Lemma 2 (Jensen).** *Let $f(z)$ be analytic throughout a disk $|z-z_0| \leq R$. Assume that $f(z_0) \neq 0$, and let $|f(z)| \leq M$ all over the disk $|z-z_0| \leq R$. Denote $n$ the number of zeroes of $f$ in the smaller disk $|z-z_0| \leq r < R$, counted with multiplicity. Then we have*

$$n\log\frac{R}{r} \leq \log\frac{1}{|f(z_0)|} + \log M. \tag{19}$$

In the theory of the Riemann zeta function $N(t)$ stands for the zeroes of $\zeta(s)$ in the rectangle with vertices 0, 1, $it$ and $1+it$; the number of zeroes in the rectangle with vertices $i(t\pm1)$ and $1+i(t\pm1)$ is thus $n(t) := N(t+1) - N(t-1)$ whenever $t \geq 1$ (and is 0 if $0 \leq t \leq 1$ according to Proposition 2, anyway). Also, it is clear that the zeroes of $\zeta(s)$ lie symmetrically with respect to the real axis, as $\zeta(x) \in \mathbb{R}$ whenever $x > 1$, and thus $\overline{\zeta(s)} = \zeta(\overline{s})$ according to the reflection principle. Our next aim is to estimate $n(t)$ at least in the order of magnitude. In the more advanced theory the Riemann–von Mangoldt formula

$$N(t) = \frac{t}{2\pi}\log\frac{t}{2\pi} - \frac{t}{2\pi} + O(\log t)$$

is proved; however, we will need only a somewhat simpler result.

**Proposition 5.** *The number of zeroes $n(t)$ of $\zeta(s)$ in the rectangle with vertices $0+i(t\pm1)$ and $1+i(t\pm1)$ does not exceed an absolute constant times $\log|t|$ for $|t| \geq \sqrt{5}-1$, and is zero if $|t| \leq \sqrt{5}-1$.*

*Proof.* Assume, as we may, $t \geq 0$. In case $t \leq \sqrt{5}-1$, the assertion is just Proposition 2. Moreover, the first rectangle fully out of the zero-free domain of Proposition 2 is $[0,1] \times [\sqrt{5}, \sqrt{5}+2]i$, so we can restrict ourselves to $t \geq \sqrt{5}+1$.

Let us apply Jensen's estimate to $\zeta(s)$ e.g. around the point $s_0 = 2 + it$ and with $R = 5/2$ and $r = \sqrt{5}$. On the one hand $|z - s_0| \leq R$ does not contain the singular point at $z = 1$, hence $\zeta$ is analytic throughout the larger disk. On the other hand the smaller disk $|z - s_0| \leq r$ covers the rectangle with vertices $0 + i(t \pm 1)$ and $1 + i(t \pm 1)$. Thus we get an estimate for $n(t)$, too. Indeed, the estimate (19) implies

$$n(t) \log \frac{\sqrt{5}}{2} \leq \log \frac{1}{|\zeta(s_0)|} + \log M \qquad \text{with} \quad M := \max_{|z-s_0|\leq 5/2} |\zeta(z)|.$$

Here by (16) $1/|\zeta(2 + it)| \leq \zeta(2) = \pi^2/6$ – an absolute constant, anyway – and Proposition 3 (iii) can be invoked to estimate $M$ as $M \ll t^{3/2}$. Whence the assertion. $\square$

We use the above estimate on the number of zeroes to derive an estimate of the order of magnitude of the *negative real part* of the logarithmic derivative of $\zeta'/\zeta$ *in the halfplane of absolute convergence.* To this end we need a well-known function theoretical lemma.

**Lemma 3 (Borel–Caratheodory).** *Let $f(z)$ be a holomorphic function in the disk $|z - z_0| \leq R$ satisfying $\Re f \leq A$ there. Then for any $r < R$ we have*

$$\max_{|z-z_0|\leq r} |f(z) - f(z_0)| \leq \frac{2r}{R - r} (A - \Re f(z_0)) . \tag{20}$$

*Proof.* It suffices to consider the case when $z_0 = 0$ and $f(0) = 0$. Consider the function

$$g(z) := \frac{f(z)}{z(2A - f(z))}.$$

Since $w := f(z)$ is in the halfplane $\Re w \leq A$, we have $|w| \leq |2A - w|$, hence on the circle $|z| = R$ we get $|g(z)| \leq 1/R$. Since $f(0) = 0$, $f(z)/z$ is regular, hence also $g$ is regular, and the maximum principle yields $|g(z)| \leq 1/R$ all over the disk $|z| \leq R$. But

$$f(z) = g(z)z(2A - f(z), \qquad f(z) = \frac{2Azg(z)}{1 + zg(z)}.$$

which shows that for $|z| = r$ the modulus of $f$ is at most

$$|f(z)| \leq \frac{2Ar/R}{1 - r/R} = \frac{2rA}{R + r}.$$

Hence the same holds throughout the disk, as was to be shown. $\square$

**Proposition 6.** *For $\sigma > 1$ and $|t| > 2$ we have*

$$\Re - \frac{\zeta'}{\zeta}(s) \leq c \log |t| - \sum_{\rho=\beta+i\gamma \ : \ \zeta(\rho)=0 \quad |t-\gamma|\leq 1} \Re \frac{1}{s - \rho}. \tag{21}$$

*Proof.* First we consider the disk $D$ around $s_0 := 2+it$ with radius 3; it lies fully in the domain $\sigma \geq -1$ and contains the rectangle $[0,1] \times i[t-1,t+1]$. In what follows let all sums, products etc. over $\rho$ stand for sums, products over all zeroes in this rectangle. By the above Proposition 5 the number $n(t)$ of terms is then $\ll \log t$.

Now put $\Phi(s) := \zeta(s)/P(s)$ with $P(s) := \prod_\rho (s-\rho)$. Then $\Phi(s)$ is regular in $D$; moreover, it is nonzero in the rectangle $[0,1] \times i[t-1,t+1]$, as well as in $\sigma > 1$, hence even its logarithm $f(s) := \log \Phi(s)$ is regular throughout the smaller disk $E := \{s \ : \ |s-s_0| \leq \sqrt{2}\}$.

The order of magnitude of $\Phi(s)$ is controlled all over the circle $\partial D$. Indeed, there $|\zeta(s)| \ll t^2 \log t \ll t^3$ in view of Proposition 3 (iv), while all the factors in $P(s)$ exceed one in absolute value, as $\partial D$ lies farther than 1 from the rectangle $[0,1] \times i[t-1,t+1]$, where all the considered zeta-roots $\rho$ lie. In all, we find

$$|\Phi(s)| \ll t^3 \qquad (s \in \partial D), \tag{22}$$

and hence all over the disk $D$ containing also $E$. Clearly, $f(s) := \log \Phi(s)$ is regular on $E$, and the Borel–Caratheodory Lemma can be applied to $f(s)$. Lemma 3 yields

$$\max_E |f(s)| \leq |f(s_0)| + \frac{2r}{R-r} \left( \max_{|s-s_0|\leq 2} \Re f(s) - \Re f(s_0) \right). \tag{23}$$

Taking into account also Proposition 5, we get

$$\Re f(s_0) \geq \log \left| \frac{\zeta(2+it)}{3^{n(t)}} \right| \geq \log \left| 1 - \sum_{n=2}^{\infty} n^{-2} \right| - c \log t \gg \log |2-\zeta(2)| - \log t \gg -\log t,$$

and also

$$\Re f(s_0) \leq \log |\zeta(2+it)| - \log 1^{n(t)} \leq \log |\zeta(2)| = c,$$

hence choosing a logarithm branch with $\Im f(s_0) = \arg \Phi(s_0) \in [-\pi, \pi)$, we obtain also $|f(s_0)| \ll \log t$. Finally, we have

$$\Re f(s) = \log |\Phi(s)| \leq \log Ct^3 \ll \log t$$

according to (22). Writing in these estimates (23) becomes

$$\max_{|s-s_0|\leq\sqrt{2}} |f(s)| \ll \frac{2r}{R-r} \left( \max_{|s-s_0|\leq 2} \Re f(s) \right) + \log t \ll \frac{R+r}{R-r} \log t \ll \log t, \tag{24}$$

in view of the choice $R=2$ and $r=\sqrt{2}$ above. Let now the point $s = \sigma + it$ be anywhere in the segment $(1+it, s_0]$. Note that for all $z$

$$\frac{\Phi'}{\Phi}(z) = \frac{\zeta'}{\zeta}(z) - \sum_\rho \frac{1}{z-\rho}. \tag{25}$$

An application of Caucy's estimate gives

$$\Re - \frac{\Phi'}{\Phi}(z) \le \left| \frac{\Phi'}{\Phi}(z) \right| = |f'(z)| \le \frac{1}{\sqrt{2}-1} \max_E |f| \ll \log t, \tag{26}$$

and combining (25) and (26) yields the assertion. □

## §3. The functional equation and the logarithmic derivative of $\zeta(s)$

In this section we use well-known properties of the $\Gamma$ function to deduce the functional equation of $\zeta(s)$. The functional equation enables us to establish various facts concerning the location of the roots of $\zeta(s)$ and to derive an even more precise formula for the logarithmic derivative.

As a generalization of the factorial function the Gamma function can be defined by its Euler integral form as

$$\Gamma(z) := \int_0^\infty x^{z-1}e^{-x}dx \qquad (\Re z > 0). \tag{27}$$

This is just the Mellin transform of the measure $e^{-x}dx/x$. Partial integration easily gives $\Gamma(z) = z\Gamma(z-1)$ and $\Gamma(1) = 1$, hence $\Gamma(n) = (n-1)!$, showing how $\Gamma$ extends the factorial function. Moreover, the very same functional equation provides a successive method to define the analytic (meromorphic) continuation of $\Gamma$ for $\Re z > -1$, then to $\Re z > -2$, etc., all over the plane. It is not too difficult to see that the integral equals to Gauss' definition

$$\lim_{m\to\infty} \frac{m!m^z}{z(z+1)\cdot\dots\cdot(z+m)}, \tag{28}$$

where the latter limit is locally uniformly convergent – hence regular – all over $\mathbb{C}\setminus -\mathbb{N}$. Gauss' definition easily yields the equivalent definition

$$\frac{1}{\Gamma(z)} = e^{Cz}z\prod_{k=1}^\infty \left(1+\frac{z}{k}\right)e^{-z/k} \qquad (C = 0.57\dots \text{ Euler's constant}) \tag{29}$$

which is the Weierstrass product representation of the entire function $1/\Gamma$. From this and the obvious Weierstrass product form of the sine function the reversal formula

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z} \tag{30}$$

obtains, too. The product form (29) clearly shows that $\Gamma$ has first order poles at points of $-\mathbb{N}$ and is regular everywhere else; moreover, $\Gamma$ is nonzero as its reciprocal is an entire function.

To establish a connection between $\zeta(s)$ and $\Gamma(z)$ we substitute $y := \lambda x$ in (27) and get

$$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx = \lambda^z\int_0^\infty y^{z-1}e^{-\lambda y}dy \qquad (\Re z > 0). \tag{31}$$

Summation of $\Gamma(z)/n^z$ over $\lambda = n \in \mathbb{N}$ gives

$$\Gamma(z)\zeta(z) = \sum_{n=1}^\infty \int_0^\infty y^{z-1}e^{-ny}dy = \int_0^\infty y^{z-1}\frac{e^{-y}}{1-e^{-y}}dy. \tag{32}$$

This leads to the functional equation through some deformation of the contour integral path, but here we follow another argument, also due to Riemann, being perhaps shorter to explain. Here instead of using $\lambda = n$ we put $\lambda = n^2$ and obtain

$$\Gamma(z)\zeta(2z) = \sum_{n=1}^{\infty} \int_0^{\infty} y^{z-1} e^{-n^2 y} dy = \int_0^{\infty} y^{z-1} \omega(y) dy \quad \left(\omega(z) := \sum_{n=1}^{\infty} e^{-n^2 y}\right). \quad (33)$$

In function theory $\omega$ is more well-known in the form of the Theta function

$$\Theta(z) := \sum_{n=-\infty}^{\infty} e^{-n^2 \pi z} \qquad \big(\Re z > 0\big), \quad (34)$$

which is related to $\omega$ as $\Theta(z) = 2\omega(\pi z) + 1$ i.e. $\omega(z) = (\Theta(z/\pi) - 1)/2$.

**Lemma 4 (Poisson–Jacobi).** *The function $\Theta(z)$ satisfies the functional equation*

$$\Theta\left(\frac{1}{z}\right) = \sqrt{z}\Theta(z) \qquad (\Re z > 0), \quad (35)$$

*with $\sqrt{z}$ denoting the principal branch of the square root function.*

*Proof.* It suffices to show the assertion for real $x > 0$, as the functional equation then extends, by analytic continuation, even to the right halfplane $\Re z > 0$.

Now consider $\tau_x(t) := e^{-\pi x t^2}$. Our point is that $\Theta(x)$ is defined as a sum, where the sum runs over function values $\tau_x(n) = e^{-\pi x n^2}$ of $\tau := \tau_x$ at $n \in \mathbb{N}$ with $x > 0$ considered a fixed parameter here. With $x > 0$ fixed, clearly $\tau \in C(\mathbb{R})$, $\tau \in L(\mathbb{R})$ and even $\tau" \in L(\mathbb{R})$. Hence Poisson's summation formula applies giving

$$\Theta(x) = \sum_{n=-\infty}^{\infty} e^{-\pi x n^2} = \sum_{n=-\infty}^{\infty} \tau(n) = \sum_{n=-\infty}^{\infty} \widehat{\tau}(n), \quad (36)$$

with the Fourier coefficients $\widehat{\tau}(n)$ defined as

$$\widehat{\tau}(n) := \int_{-\infty}^{\infty} e^{-2\pi i n t} \tau(t) dt = \int_{-\infty}^{\infty} e^{-2\pi i n t - \pi x t^2} dt = e^{-\pi n^2/x} \int_{-\infty}^{\infty} e^{-\pi x (t + in/x)^2} dt. \quad (37)$$

Here the last integral is a complex path integral over the real line, but can be moved to another horizontal straight line passing through $-in/x$. The integrand $e^{-\pi x(t+in/x)^2}$ is an entire function of $t \in \mathbb{C}$, hence no residues occur, while integrals over the segments $[T, T - in/x]$ in the deformation process contribute o(1) when $T \to \pm\infty$. In all,

$$\int_{-\infty}^{\infty} e^{-\pi x(t+in/x)^2} dt = \int_{-\infty - in/x}^{\infty - in/x} e^{-\pi x (z + in/x)^2} dz = \int_{-\infty}^{\infty} e^{-\pi x u^2} du = \frac{1}{\sqrt{x}}, \quad (38)$$

computing the integral over $z$ using parametrization by $z = u - ix/n$ with $u \in \mathbb{R}$. Taking into account (34), collecting (36), (37) and (38) furnishes (35). $\square$

**Proposition 7.** *The Riemann zeta function $\zeta(s)$ is analytic all over $\mathbb{C}$ except for a simple pole of residuum 1 at $s = 1$. Moreover, it admits the functional equation* (4).

*Proof.* Setting $s = z/2$ with $\Re s > 1$ in (33) and inserting $\omega(y) = (\Theta(y/\pi) - 1)/2$, substituting $x := y/\pi$ gives

$$\Gamma\left(\frac{s}{2}\right)\zeta(s) = \int_0^\infty y^{s/2-1}\frac{\Theta(y/\pi)-1}{2}dy = \pi^{s/2}\int_0^\infty x^{s/2-1}\frac{\Theta(x)-1}{2}dx = \int_0^1 + \int_1^\infty . \tag{39}$$

After dividing by $\pi^{s/2}$, an application of Lemma 4 in the integral over $[0,1]$ leads to

$$\frac{2}{\pi^{s/2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \int_0^1 x^{s/2-1}\left(\sqrt{1/x}\Theta(1/x) - 1\right)dx + 2\int_1^\infty x^{s/2-1}\omega(\pi x)dx \quad \left(\Re s > 1\right).$$

Substituting $u = 1/x$ in the first integral then yields

$$\begin{aligned}
\int_0^1 x^{s/2-1}\left(\sqrt{1/x}\Theta(1/x) - 1\right)dx &= \int_0^1 x^{s/2-1}\left(x^{-1/2}\left(\Theta(1/x)-1\right) + x^{-1/2} - 1\right)dx \\
&= \int_1^\infty u^{-1/2-s/2}\left(\Theta(u)-1\right)du + \int_1^\infty \left(u^{-1/2-s/2} - u^{-1-s/2}\right)du \\
&= 2\int_1^\infty u^{-1/2-s/2}\omega(\pi u)du + \left(\left[\frac{2u^{1/2-s/2}}{1-s}\right]_1^\infty - \left[\frac{2u^{-s/2}}{-s}\right]_1^\infty\right) \\
&= 2\int_1^\infty x^{-1/2-s/2}\omega(\pi x)dx + \frac{2}{s-1} - \frac{2}{s} \\
&= 2\int_1^\infty x^{(1-s)/2-1}\omega(\pi x)dx + \frac{2}{s(s-1)} \qquad \left(\Re s > 1\right).
\end{aligned}$$

and thus

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \int_1^\infty \left(x^{(1-s)/2-1} + x^{s/2-1}\right)\omega(\pi x)dx + \frac{1}{s(s-1)} \quad \left(\Re s > 1\right). \tag{40}$$

Now if we multiply by $s(s-1)/2$, taking into account $\Gamma(s/2)s/2 = \Gamma(s/2+1)$ we obtain

$$\xi(s) = (s-1)\pi^{-s/2}\Gamma\left(\frac{s}{2}+1\right)\zeta(s) = P(s) + P(1-s) + 2 \quad \left(\Re s > 1\right), \tag{41}$$

where

$$P(z) := \frac{z(z-1)}{2}\int_1^\infty x^{z/2-1}\omega(\pi x)dx.$$

Thus even if originally we have started with $\Re s > 1$, the end formula gives an entire function, since $P(z)$ is regular all over $\mathbb{C}$. Moreover, it is easy to see that substituting

$s$ by $1-s$ the formula in (41) for $\xi(s)$ remains unchanged; thus we arrive at (4). Expressing now $\zeta(s)$ gives

$$\zeta(s) = \frac{1}{s-1}\pi^{s/2}\frac{1}{\Gamma\left(\frac{s}{2}+1\right)}\xi(s), \tag{42}$$

all terms on the right hand side being entire functions but for $1/(s-1)$. As for $s=1$ we already know the behavior of $\zeta(s)$, the meromorphic behavior of $\zeta(s)$ has now been fully described. Finally, observe that $\xi(s) \neq 0$ for $\Re s > 1$, as neither terms can vanish in the definition (4) of $\xi$ if $\Re s > 1$. It follows by the functional equation just proved, that $\xi$ does not vanish for $\Re s < 0$ either. Recalling (42) it follows that for $\Re s < 0$ $\zeta(s)$ vanishes exactly when $1/\Gamma(s/2+1)$, that is, when $s \in -2\mathbb{N}$. $\square$

The zeroes $s = -2, -4, \ldots$ are called the *trivial zeroes* of $\zeta(s)$, while all other zeroes – the so-called *nontrivial zeroes* – lie in the *critical strip* $0 \leq \Re s \leq 1$. Clearly, any *nontrivial* zero $\rho$ of $\zeta$ is a zero of $\xi$, too; and conversely, a trivial zero of $\zeta$ is never a zero of $\xi$, for the effect of the corresponding pole of the $\Gamma$ function. By the functional equation it follows that for a nontrivial root $\rho$ $\zeta(\rho) = 0$ implies $\zeta(1-\rho) = 0$; also, by the reflection principle we see that $\overline{\rho}$ and $1-\overline{\rho}$ are $\zeta$-roots, too.

Hadamard's approach was then to continue with the build-up of general complex function theory, in particular of entire functions, to see that there are infinitely many nontrivial zeroes lying in the critical strip. The best way to see this is to observe, that by the functional equation $\Xi(z) := \xi(z+1/2)$ is symmetric with respect to both the real and the imaginary axis; hence it is a function of $z^2$ only (Exercise!). Then taking $\Xi(z) = \Phi(z^2)$, the function $\Phi(w)$ is another entire function, but its growth order is now $1/2$. (It is not too difficult to see that the order of $\xi$ is the order of $\Gamma$, which is 1.) However, general function theory says that an entire function of order $< 1$ is the Weierstrass product of its root factors; in particular, if the function is not a polynomial, then it must have infinitely many zeroes.

In any case, the entire function $\xi$ has all its zeroes in the critical strip. Moreover, it has order 1, since taking into account (41), it suffices to see that $P(z)$ is of order 1, which is easy to calculate. (In fact, we will need only that $\xi$ has order *at most* 1, which is easier to see from upper estimate of the growth of $P(z)$.) In all, we find the Weierstarss product representation

$$\xi(s) = e^{as+b} \prod_{\rho \,:\, \xi(\rho)=0} \left(1 - \frac{s}{\rho}\right) e^{s/\rho}, \tag{43}$$

which gives

$$\frac{\xi'}{\xi}(s) = a + \sum_{\rho \,:\, \xi(\rho)=0} \left(\frac{1}{s-\rho} + \frac{1}{\rho}\right), \tag{44}$$

which is a meromorphic representation by Mittag-Leffler sums. Writing in the definition of $\xi$ we finally obtain

$$\frac{\zeta'}{\zeta}(s) = -\frac{1}{s-1} + \frac{1}{2}\log\pi - \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}+1\right) + a + \sum_{\rho\,:\,\xi(\rho)=0}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right). \qquad (45)$$

**Lemma 5.** *The logarithmic derivative of* $\Gamma(z)$ *satisfies*

$$\frac{\Gamma'}{\Gamma}(z) = \log z + O(1) \qquad (\Re z > 0, |z| > 1).$$

*Proof.* The Mittag-Leffler partial fractions sum formula for the logarithmic derivative of $\Gamma$ follows from the product representation (29) as

$$-\frac{\Gamma'}{\Gamma}(z) = C + \frac{1}{z} + \sum_{k=1}^{\infty}\left(\frac{1}{k+z}-\frac{1}{k}\right) = C + \frac{1}{z} - \sum_{k=1}^{\infty}\left(\frac{z}{(k+z)k}\right). \qquad (46)$$

Now we write

$$\sum_{k=1}^{N}\left(\frac{1}{k+z}-\frac{1}{k}\right) \approx \int_1^N\left(\frac{1}{x+z}-\frac{1}{x}\right)dx = \log(N+z) - \log(z+1) - \log N \approx \log z,$$

where it is easy to see that the deviation of the expressions connected by the approximate equality symbol $\approx$ will remain bounded in $\Re z > 0$. □

Using Lemma 5 in (45) furnishes the formula

$$\frac{\zeta'}{\zeta}(s) = -\frac{1}{2}\log\left(\frac{s}{2}+1\right) + \sum_{\rho}^{*}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right) + O(1) \quad (s\to\infty \text{ in } \Re s > 0), \qquad (47)$$

with $\sum_\rho^*$ denoting a sum extended over the nontrivial roots of $\zeta(s)$. Since for $\Re s > 1$ all terms in the sum have positive real parts, we finally obtain the next assertion.

**Proposition 8.** *We have*

$$\Re - \frac{\zeta'}{\zeta}(s) \le C + \frac{1}{2}\log t - \sum_{\rho}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right) \quad (s = \sigma + it,\ \sigma > 1,\ |t| \ge 2), \qquad (48)$$

*with* $\sum_\rho$ *being extended over all, or over any subset of the nontrivial roots of* $\zeta(s)$.

## §4. Basic results on the error term in prime distribution

**Lemma 6.** *Let $\delta(x)$ be any monotone increasing function satisfying $\sqrt{x} \le \delta(x) \le x$. If $\Psi(x) = x + O(\delta(x))$, then $\pi(x) = \mathrm{li}(x) + O(\delta(x))$.*

*Proof.* Recall that $\Psi(x) := \sum_{n<x} \Lambda(n)$ with $\Lambda$ being von Mangoldt's function

$$\Lambda(n) := \begin{cases} \log p & \text{if} \quad n = p^k \quad \text{with} \quad p \quad \text{prime, and} \quad k \in \mathbb{N} \\ 0 & \text{otherwise.} \end{cases}$$

Let $l(n) := \log p$ for $n = p$ prime and 0 otherwise; then by definition $\Theta(x) := \sum_{n<x} l(n)$. First we compare $\Psi(x)$ and $\Theta(x)$, which is easy:

$$\begin{aligned} \Theta(x) \le \Psi(x) = \sum_{p<x} \sum_{k\,:\,p^k \le x} \log p &= \Theta(x) + \sum_{p} \sum_{2 \le k \le \frac{\log x}{\log p}} \log p \\ &< \Theta(x) + \sum_{p \le \sqrt{x}} \left[\frac{\log x}{\log p}\right] \log p \le \Theta(x) + \pi(\sqrt{x}) \log x = \Theta(x) + O(\sqrt{x}), \end{aligned}$$

using also Chebyshev's theorem $\pi(y) \ll y/\log y$. So it follows that $\Psi(x) - \Theta(x) = O(\sqrt{x}) = O(\delta(x))$, and it suffices to compare $\Theta(x) - x$ and $\pi(x) - \mathrm{li}(x)$.

We apply partial summation, then the assumption, and bring back the main term by a reversed application of partial integration. Hence

$$\begin{aligned} \pi(x) = \sum_{n \le x} \frac{l(n)}{\log n} &= \Theta(x) \frac{1}{\log x} - \int_2^x \Theta(y) \frac{-1}{y \log^2 y} dy \\ &= \frac{x}{\log x} + \frac{O(\delta(x))}{\log x} + \int_2^x (y + O(\delta(y))) \frac{dy}{y \log^2 y} \\ &= \frac{x}{\log x} + \int_2^x \frac{y}{y \log^2 y} dy + O(\delta(x)) + O(\delta(x)) \int_2^x \frac{dy}{y \log^2 y} \\ &= \int_2^x \frac{dy}{\log y} + O(\delta(x)) = \mathrm{li}(x) + O(\delta(x)). \end{aligned}$$

The assertion obtains. □

**Lemma 7.** *Let $\delta(x)$ be any monotone increasing function with $\sqrt{x} \le \delta(x) \le x$. If*

$$T(x) := \sideset{}{^*}\sum_{n<x} \Lambda(n) \log\left(\frac{x}{n}\right) = x + O(\delta(x)), \tag{49}$$

*then $\pi(x) = \mathrm{li}(x) + O(\delta(x))$.*

*Proof.* Let us define $f(n) := \Lambda(n)$ and $g(y) := \log(x/y) = \log x - \log y$, where here $x$ is considered a fix constant, and $y$ is the variable between 1 and $x$. We apply partial summation to get

$$\Psi(x) = -\int_1^x T(y)\frac{-1}{y}dy.$$

Here both integrated terms vanish as $T(1) = 0 = \log(x/x)$. Observe that $T(y) = 0$ for all $y < 2$, hence the lower limit of the integration can be changed to 2. Now we get

$$\Psi(x) = \int_2^x (y + O(\delta(y))\frac{1}{y}dy = x + O\left(\int_2^x \frac{\delta(y)}{y}dy\right) = x + O(\Delta(x)),$$

with $\Delta(x) := \int_2^x \frac{\delta(y)}{y}dy$. Repeating the argument of Lemma 6 with $\Delta$, we get

$$\pi(x) = \mathrm{li}(x) + \frac{O(\Delta(x))}{\log x} + \int_2^x \frac{O(\Delta(y))}{y\log^2 y}dy,$$

hence

$$\pi(x) - \mathrm{li}(x) = O\left(\frac{1}{\log x}\int_2^x \frac{\delta(y)}{y}dy + \int_2^x\int_2^y \frac{\delta(u)}{u}du\frac{dy}{y\log^2 y}\right).$$

Clearly, the first term is $O(\delta(x))$. Interchanging the integrals in the second we find

$$\begin{aligned}\int_2^x\int_2^y \frac{\delta(u)}{u}du\frac{dy}{y\log^2 y} &= \int_2^x \frac{\delta(u)}{u}\int_u^x \frac{dy}{y\log^2 y}du = \int_2^x \frac{\delta(u)}{u}\left(\frac{1}{\log u} - \frac{1}{\log x}\right) \\ &= \int_2^{\sqrt{x}} + \int_{\sqrt{x}}^x \le \sqrt{x} + \frac{1}{\log x}\int_2^x \frac{\delta(u)}{u},\end{aligned}$$

which is exactly the same as before apart from the term $\sqrt{x} \le \delta(x)$. Hence we get $O(\delta(x))$, as stated. □

## §5. Mellin transform and Perron's coefficient formula

In this section we assume that $a_n$ is a sequence (an arithmetical function), and consider the corresponding Dirichlet series

$$A(s) := \sum_{n=1}^{\infty} \frac{a_n}{n^s}, \tag{50}$$

where the series will be assumed to be convergent. Putting $\alpha(x) := \sum_{n<x} a_n$, the series (50) (if convergent) is just

$$A(s) = \int_1^{\infty} x^{-s} d\alpha(x), \tag{51}$$

which is the general form of a Stieltjes–Mellin transform whenever $\alpha$ is a function of bounded variation, or is just a measure. Substituting $t := \log x$, we obtain

$$A(s) = \int_0^{\infty} e^{-ts} d\alpha(e^t), \tag{52}$$

which is just the Laplace transform of the measure $d\beta(t) := d\alpha(e^t)$. With a change of variable this can be transformed to a Fourier transform as well; however, it is already clear, that similar inversion formulae, as for e.g. the Laplace transform, should hold. One example is to follow after the next technical lemma.

**Lemma 8.** *Let $r > 0$. Then – interpreting the integral in the Cauchy principal value sense – we have*

$$\int_{r-i\infty}^{r+i\infty} \frac{y^s}{s} ds = I(y) := \begin{cases} 2\pi i & \text{if} \quad y > 1 \\ \pi i & \text{if} \quad y = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{53}$$

*Moreover, let $T > 0$ be any parameter. Then we have*

$$\left| \int_{r-iT}^{r+iT} \frac{ds}{s} - \pi i \right| \le \frac{2r}{T} \qquad \text{and} \qquad \left| \int_{r-iT}^{r+iT} \frac{y^s}{s} ds - I(y) \right| \le \frac{2y^r}{T|\log y|}. \tag{54}$$

*Proof.* It suffices to prove the second assertion. Let first define $\eta := \log y$; then $\eta$ is above or below 0 when $y$ is above or below 1. In case $y = 1$, the integral is

$$\int_{r-iT}^{r+iT} \frac{ds}{s} = \int_0^T \left( \frac{i}{r+it} + \frac{i}{r-it} \right) dt = 2i \int_0^T \frac{r}{r^2+t^2} dt = 2i \arctan(T/r).$$

Since $\pi/2 - \arctan x = \arctan(1/x) < 1/x$, we obtain the result if $y = 1$.

Let now $y < 1$, i.e. $\eta < 0$. We move the path of integration to the right, making use of the fact that the full circle integral along the boundary of the rectangle $Q$ with vertices $r - iT$,$r + iT$, $R + iT$ and $R - iT$, vanishes by Cauchy's Fundamental Theorem since $y^s/s$ is regular on $Q$. On the other hand the integral along the right side is $O(2T \exp(R\eta)/R)$, which tends to 0 rapidly together with $R$ because $\eta < 0$. This makes it possible to shift the curve of integration to a disconnected pair of horizontal lines from $r \pm iT$ to $\infty \pm iT$. On these segments we can estimate the integrals as

$$\left|\int_{r\pm iT}^{\infty\pm iT} \frac{\exp(\eta s)}{s} ds\right| \leq \int_{r}^{\infty} \frac{\exp(\eta x)}{T} dx = \frac{y^r}{|\eta| T}.$$

Finally, let $y > 1$. Then the path of integration is moved towards the left. First we push the path of integration to the rectangle with vertices $r \pm iT$, $-R \pm iT$, and then, checking that the integral on the leftmost line $[-R - it, -R + iT]$ is $O(y^{-R}T/R) \to 0$, we even transform the integral to the two lines $[r \pm iT, -\infty \pm iT]$. As before, both line integrals can be estimated as

$$\left|\int_{r\pm iT}^{-\infty\pm iT} \frac{\exp(\eta s)}{s} ds\right| \leq \int_{-\infty}^{r} \frac{\exp(\eta x)}{T} dx \leq \frac{y^r}{\eta T}.$$

The only difference is that here the integrand is not analytic, but has one simple pole at $s = 0$. Hence by the residuum theorem the total line integral over the closed curve is not 0, but is $2\pi i \mathrm{Res}[y^s/s]_{s=0}$. As the residuum is 1, the assertion obtains. $\square$

**Lemma 9 (Perron).** *Let $r > 0$ and assume that the Dirichlet series $A(s) := \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ converges absolutely and uniformly for $\sigma \geq r$. Then – interpreting the integral in the Cauchy principal value sense – we have*

$$\frac{1}{2\pi i} \int_{r-i\infty}^{r+i\infty} A(s) \frac{x^s}{s} ds = \sideset{}{^*}\sum_{n<x} a_n, \tag{55}$$

*where the * denotes the interpretation that in case $x = m \in \mathbb{N}$ the corresponding term is taken with weight $1/2$. Moreover, let $T > 0$ be any parameter. Then we have*

$$\left| \frac{1}{2\pi i} \int_{r-iT}^{r+iT} A(s) \frac{x^s}{s} ds - \sideset{}{^*}\sum_{n<x} a_n \right| \leq \frac{x^r}{\pi T} \sum_{n=1}^{\infty} \frac{|a_n|}{|\log^*(x/n)| n^r}, \tag{56}$$

*where here $\log^*(y)$ denotes $\log(y)$ for $0 < y \neq 1$ and $1$ for $y = 1$.*

*Proof.* By uniform convergence, we can split the Dirichlet series to terms and apply the preceding Lemma 8 to each term with $y = x/n$. The first part of the assertion then follows. Moreover, in the finite interval case the error terms add up as

$$\sum_{n=1}^{\infty} |a_n| \frac{2(x/n)^r}{|\log^*(x/n)|T} = \frac{2x^r}{T} \sum_{n=1}^{\infty} \frac{|a_n|}{|\log^*(x/n)|n^r}.$$

□

The reader will find no difficulty in expressing the individual terms $a_n$ from the coefficient formulae for $x = n + 1/2$ and $x = n - 1/2$, e.g.

## §6. Another contour integral deduction

There are many variants of the above connections between the transform and the original function or measure. We will directly employ the following version.

**Lemma 10.** *Let $r > 0$. Then we have*

$$\int_{r-i\infty}^{r+i\infty} \frac{y^s}{s^2} ds = L(y) := \begin{cases} 2\pi i \log y & \text{if} \quad y \geq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{57}$$

*Moreover, let $T > 0$ be any parameter. Then we have*

$$\left| \int_{r-iT}^{r+iT} \frac{y^s}{s^2} ds - L(y) \right| \leq \frac{2\pi y^r}{T}. \tag{58}$$

*Proof.* It suffices to prove the second assertion. Put $\eta := \log y$. In case $y = 1$, the integral is

$$\int_{r-iT}^{r+iT} \frac{ds}{s^2} = \left[ \frac{-1}{s} \right]_{r-iT}^{r+iT} = \frac{1}{r-iT} - \frac{1}{r+iT} = \frac{2iT}{r^2+T^2},$$

which is less than $2/T$ in absolute value. Whence the result if $y = 1$.

Let now $y < 1$, i.e. $\eta < 0$. As before, we apply Cauchy's Fundamental Theorem along the boundary of the rectangle $Q$ with vertices $r-iT$,$r+iT$, $R+iT$ and $R-iT$: the full circle integral vanishes since $y^s/s^2$ is regular on $Q$. On the other hand the integral along the right side is $O(2T\exp(R\eta)/R^2)$, which tends to 0 rapidly together with $R$ because $\eta < 0$. This makes it possible to shift the curve of integration to a disconnected pair of horizontal lines from $r \pm iT$ to $\infty \pm iT$. On these segments we can estimate the integrals as

$$\left| \int_{r\pm iT}^{\infty \pm iT} \frac{\exp(\eta s)}{s^2} ds \right| \leq \int_r^\infty \frac{\exp(\eta x)}{T^2+x^2} dx < \frac{y^r}{T} \int_0^\infty \frac{T}{T^2+x^2} dx = \frac{\pi y^r}{2T}.$$

Finally, let $y > 1$. Then the path of integration is moved towards the left. First we push the path of integration to the rectangle with vertices $r \pm iT$, $-R \pm iT$, and then, checking that the integral on the leftmost line $[-R - it, -R + iT]$ is $O(y^{-R}T/R^2) \to 0$, we even transform the integral to the two lines $[r \pm iT, -\infty \pm iT]$. Similarly as before, but using also $0 < r < T$, both line integrals can be estimated as

$$\left| \int_{r\pm iT}^{-\infty \pm iT} \frac{\exp(\eta s)}{s^2} ds \right| \leq \int_{-\infty}^r \frac{\exp(\eta x)}{T^2+x^2} dx < \frac{y^r}{T} \int_{-\infty}^\infty \frac{T}{T^2+x^2} dx = \frac{\pi y^r}{T}.$$

The only difference is that here the integrand is not analytic, but has one simple pole at $s = 0$. Hence by the residuum theorem the total line integral over the closed curve is not 0, but is $2\pi i \mathrm{Res}[y^s/s^2]_{s=0}$. Clearly, this residuum is $\eta = \log y$. Hence the assertion. $\square$

**Lemma 11 (Perron).** *Let $r > 0$ and assume that the Dirichlet series $A(s) := \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ converges absolutely and uniformly for $\sigma \geq r$. Then – interpreting the integral in the Cauchy principal value sense – we have*

$$\int_{r-i\infty}^{r+i\infty} A(s)\frac{x^s}{s^2}ds = 2\pi i \sum_{n<x} a_n \log\left(\frac{x}{n}\right). \tag{59}$$

*Moreover, let $T > 0$ be any parameter. Then we have*

$$\left| \frac{1}{2\pi i}\int_{r-iT}^{r+iT} A(s)\frac{x^s}{s^2}ds - \sum_{n<x} a_n \log\left(\frac{x}{n}\right)\right| \leq \frac{x^r}{T}\sum_{n=1}^{\infty}\frac{|a_n|}{n^r}. \tag{60}$$

*Proof.* By uniform convergence, we can split the Dirichlet series to terms and apply the preceding Lemma 8 to each term with $y = x/n$. The first part of the assertion then follows. Moreover, in the finite interval case the error terms add up as

$$\sum_{n=1}^{\infty} |a_n| \frac{(x/n)^r}{T} = \frac{x^r}{T}\sum_{n=1}^{\infty}\frac{|a_n|}{n^r}.$$

□

## §7. Connection of $\zeta(s)$ and $\pi(x)$

The following are the key results of Landau's analysis – apart from the loss in error precision of a factor 2 under the exponent, that is, a square root – in case of Theorem 2, due to the weaker form of Lemma 7 used by him.

**Theorem 1 (Landau).** *Assume that the Riemann zeta function satisfies with some constants $A, C, a > 0$ and $d \geq 1$*

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right| \leq C \log^A |t| \qquad |t| \geq 2,\, \sigma \geq 1 - \frac{a}{\log^d |t|}. \tag{61}$$

*Then we have with any $B > d+1$ the inequality*

$$\pi(x) = \mathrm{li}(x) + O(xe^{-x^{1/B}}). \tag{62}$$

Moreover,

**Theorem 2 (Pintz).** *Assume that the Riemann zeta function satisfies with some constant $a > 0$*

$$\zeta(s) \neq 0 \qquad |t| \geq 2,\, \sigma \geq 1 - \frac{a}{\log |t|}. \tag{63}$$

*Then we have with any $b < 2\sqrt{a}$ the inequality*

$$\pi(x) = \mathrm{li}(x) + O(xe^{-b\sqrt{x}}). \tag{64}$$

The proof of Theorem 2 needs a further proposition, that of obtaining an upper estimate of $\zeta'/\zeta$ from knowing a merely zero-free region.

**Proposition 9.** *Assume that the Riemann zeta function satisfies with some constants $a, d > 0$*

$$\zeta(s) \neq 0 \qquad |t| \geq 2,\, \sigma \geq 1 - \frac{a}{\log^d |t|}. \tag{65}$$

*Then for any positive $\epsilon > 0$ we have*

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right| \leq C \log^{1+2d} |t| \qquad |t| \geq 2,\, \sigma \geq 1 - \frac{a-\epsilon}{\log^d |t|}, \tag{66}$$

*with $C = C(\epsilon)$ being a suitable constant.*

*Proof.* The statement is obvious for $\sigma \geq 2$, so consider only $\sigma < 2$. Assume, as we may, $t > 0$, and consider the domain

$$\mathcal{D} := \{s \in \mathbb{C} \ :\ s = \sigma + it,\ \sigma \geq 1 - a\log^{-d} t,\ t \geq 1\}.$$

In this simply connected domain $\zeta(s)$ is regular and is nonzero by condition and in view of Proposition 2. Hence its logarithm $\log \zeta(s)$, and also the logarithmic derivative $(\log \zeta)' = \zeta'/\zeta$ are analytic in $\mathcal{D}$. It follows that fixing any finite value of $t_0$, the assertion holds true for $2 \leq t \leq t_0$ with some suitable constant $C$. So we assume also $t > t_0$ with $t_0$ chosen so that $a \log^{-d} t < 1/4$ and that $(a - \epsilon/3) \log^d (t+2) < a \log^d t$.

Let now $s = \sigma + it$ be arbitrary in $\mathcal{D}$ with $t > t_0$ and consider the point $s_0 := 2+it$. The circle with radius $R := 1 + (a - \epsilon/3) \log^{-d} t$ around $s_0$ lies fully within the domain $\mathcal{D}$ by condition and the choice of $t_0$. In this domain Proposition 3 provides $\log |\zeta(s)| \leq 1/2 \log t + c$, since for $1 - \sigma \leq 1/\log t$ part (i), otherwise (iii) can be invoked.

Next we will apply the Borel–Caratheodory Lemma (i.e. Lemma 3). In our application of this, we choose $f := \log \zeta$, $z_0 = s_0 = 2 + it$, $R = 1 + (a - \epsilon/3) \log^{-d} t$ and $r = R - \epsilon/3 \log^{-d} t$. We conclude that in the circle $|z - s_0| \leq r$ the estimate

$$|\log \zeta(z) - \log \zeta(2+it)| \leq \frac{2r(\log(t+2) + C - \Re \log \zeta(2+it))}{R-r} \ll \frac{\log^{d+1} t}{\epsilon}$$

holds. Hence also $|\log \zeta(z)| \ll \log^{d+1} t/\epsilon$. Take now an $s = \sigma + it$ with $t > t_0$ and $\sigma \geq 1 - (a - \epsilon) \log^{-d} t$. Note that the small disk $|z - s| \leq \rho$ with radius $\rho := (\epsilon/3) \log^{-d} t$ at $s$ lies fully within the previous disk $|z - s_0| \leq r$. By the Cauchy integral formula we find

$$\frac{\zeta'(s)}{\zeta(s)} = \frac{1}{2\pi} \int_0^{2\pi} \frac{\log \zeta(s + \rho e^{i\varphi})}{\rho e^{i\varphi}} d\varphi,$$

hence

$$\left| \frac{\zeta'(s)}{\zeta(s)} \right| \leq \frac{c \log^{d+1} t}{\epsilon \rho} \ll \frac{\log^{2d+1} t}{\epsilon^2}$$

□

Having Proposition 9, the proof of Theorem 1 and 2 are contained in the following combined result.

**Theorem 3.** *Assume that the Riemann zeta function satisfies with some constants* $A, C, a > 0$ *and* $d \geq 1$

$$\left| \frac{\zeta'(s)}{\zeta(s)} \right| \leq C \log^A |t| \qquad |t| \geq 2, \, \sigma \geq 1 - \frac{a}{\log^d |t|}. \tag{67}$$

*Set* $D := (d+1) d^{\frac{1}{d+1}}/d$. *Then we have with any* $b < a^{1/(d+1)} D$ *the inequality*

$$\pi(x) = \mathrm{li}(x) + O(x e^{-b \log^{\frac{1}{d+1}} x}). \tag{68}$$

*Proof.* The proof hinges upon an application of the residuum theorem on a relatively complicated contour integration path. Let us choose $x > 2$ and $T > 2$. Recalling the definition of $T(x)$ from (49), we start with the formula

$$\left| T(x) - \frac{1}{2\pi i} \int_{2-iT}^{2+iT} \left( -\frac{\zeta'(s)}{\zeta(s)} \right) \frac{x^s}{s^2} ds \right| \le \frac{x^2}{T} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^2} = C \frac{x^2}{T}, \tag{69}$$

following directly from Lemma 11 and the fact that $\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^2} = -\zeta'(2)/\zeta(2)$ is just a constant. Now we shift the line of integration to a new path $\Gamma$ from $2 - iT$ to $2+iT$, chosen so that joining $-\Gamma$ and the vertical straight line segment $[2-iT, 2+iT]$ forms a closed Jordan curve, encircling the pole of $\zeta$ at $s = 1$. Hence if we change the line of integration to the new path $\Gamma$, by the residuum theorem we obtain the same integral least $2\pi i$ the value of the residuum at 1. That is, we find

$$\int_{[2-iT,2+iT]} \left( -\frac{\zeta'(s)}{\zeta(s)} \right) \frac{x^s}{s^2} ds = \int_{\Gamma} \left( -\frac{\zeta'(s)}{\zeta(s)} \right) \frac{x^s}{s^2} ds + 2\pi i x. \tag{70}$$

In the following we estimate the contribution of the line integral along $\Gamma$. $\Gamma$ will be chosen symmetric to the real axis, i.e. $\Gamma = \Gamma_+ \bigcup \Gamma_-$ with $\Gamma_-$ joining $2 - iT$ and $1/2$ within the halfplane $\Im z \le 0$, and $\Gamma_+ := \overline{\Gamma}_-$ joining $1/2$ and $2 + iT$ within the upper halfplane $\Im z \ge 0$. Clearly it suffices to describe both the definition and the estimates of the various parts of $\Gamma_+$. Now we define

$$\Gamma_+ := \Gamma_0 \cup \gamma \cup \Gamma^* \quad \text{with} \quad \Gamma_0 := \left[ \frac{1}{2}, \frac{1}{2} + 2i \right], \quad \Gamma^* := \left[ 1 - \frac{a}{\log^d T} + iT, 2 + iT \right]$$
$$\text{and} \qquad \gamma(t) := 1 - \frac{a}{\log^d t} + it \quad \text{for} \quad t \in [2, T]. \tag{71}$$

On the line segment $\Gamma_0$ the Riemann zeta function is analytic, and also nonzero by Proposition 2, hence its logarithmic derivative is continuous and thus bounded. (In fact, a lower estimate of $\zeta(s)$ is already at hand, and a similar upper estimate or an estimate using Cauchy's integral formula would be easy to prove.) We thus find

$$\left| \int_{\Gamma_0} \left( -\frac{\zeta'(s)}{\zeta(s)} \right) \frac{x^s}{s^2} ds \right| \le 2 \frac{C\sqrt{x}}{1/4} = O(\sqrt{x}). \tag{72}$$

When $t \ge 2$, the estimate of the logarithmic derivative of $\zeta(s)$ uses the assumption (61). Similarly as above,

$$\left| \int_{\Gamma^*} \left( -\frac{\zeta'(s)}{\zeta(s)} \right) \frac{x^s}{s^2} ds \right| \le 2C \log^A T \frac{x^2}{T^2} = O(\frac{x^2}{T}), \tag{73}$$

since $\log^A T = O(T)$. Finally we return to the integral over the essential portion $\gamma$. Here by (61) we can write

$$\left|\int_\gamma \left(-\frac{\zeta'(s)}{\zeta(s)}\right)\frac{x^s}{s^2}ds\right| \le C\log^A T\int_2^T \frac{x^{1-a/\log^d t}}{t^2}\left|\frac{ad}{t\log^{d+1}t}+1\right|dt$$
$$\ll x\log^A T\int_2^T \frac{x^{-a/\log^d t}}{t^2}dt. \tag{74}$$

It remains to estimate the latter integral. Put $\mu := a\log x$, extend the interval of integration to $[1,\infty)$, and apply a change of variable $v := \log t$ and then $u := \mu^{-1/(d+1)}v$ to get

$$\int_2^T \frac{x^{-a/\log^d t}}{t^2}dt < \int_0^\infty \exp\left(-\frac{\mu}{v^d}-v\right)dv = \nu\int_0^\infty \exp\left(\nu\left\{-\frac{1}{u^d}-u\right\}\right)du, \tag{75}$$

where $\nu := \mu^{1/(d+1)}$. The last integral is of the form

$$\int_0^\infty \exp(\nu h(u))du,$$

where $h(u) = -u^{-d}-u$ is a strictly concave function attaining its maximum at one single inner point $\alpha := d^{1/(d+1)}$ of the interval. Also, it is several times continuously differentiable (in fact, analytic) along the interval. (Around 0 we may cut a fixed, small interval to get rid of the singularity where $h(+0) = -\infty$; on this interval any rough estimate will do, being much less then the contribution around $\alpha$.) We can thus apply

**Lemma 12 (Laplace).** *Let $\Phi(x)$ and $h(x)$ be two real continuous functions in the finite or semi-infinite interval $\alpha \le x \le \beta$, such that*

i) *$\Phi(x)e^{\nu h(x)}$ is absolutely integrable over the interval for every positive value of $\nu$;*

ii) *$h(x)$ has a single maximum in the interval, namely at $x = \alpha$; and the supremum of $h(x)$ in any closed subinterval not containing $\alpha$ is less than $h(\alpha)$;*

iii) *$h"(x)$ is continuous; and $h'(\alpha) = 0$, $h"(\alpha) < 0$.*

*Then, as $\nu \to \infty$, we have*

$$\int_\alpha^\beta \Phi(x)e^{\nu h(x)}dx \sim \Phi(\alpha)e^{\nu h(\alpha)}\left\{\frac{-\pi}{2\nu h"(\alpha)}\right\}^{1/2}. \tag{76}$$

*Proof.* This is the so called Laplace method, see e.g. [7, §18, p. 39]. □

In fact, we have $\Phi \equiv 1$, and shall apply the asymptotic expansion (76) both for $[\alpha, \infty)$ and for $(0, \alpha]$. So we get

$$\int_0^\infty \exp\left(\nu\left\{-\frac{1}{u^d} - u\right\}\right) du \sim 2e^{\nu h(\alpha)} \left\{\frac{-\pi}{2\nu h"(\alpha)}\right\}^{1/2}.$$

Here $h(\alpha) = -d^{-d/(d+1)} - d^{1/(d+1)} = -d^{1/(d+1)}(d+1)/d$ and $h"(\alpha) = -(d+1)/d^{1/(d+1)}$, hence

$$\nu \int_0^\infty \exp\left(\nu\left\{-\frac{1}{u^d} - u\right\}\right) du \sim 2e^{-\nu d^{1/(d+1)}(d+1)/d} \left\{\frac{\pi d^{1/(d+1)}\nu}{2(d+1)}\right\}^{1/2}. \tag{77}$$

**Remark.** Although it is good to see that our estimate is asymptotically the right value, deriving an essentially sharp upper estimate suffices. In what follows we show such an estimate of the last integral of (75).

*Proof.* Write $D := d^{1/(d+1)}(d+1)/d$; it is a relevant constant being the minimum value of $\varphi(u) := u^{-d} + u$ on $(0, \infty)$. Indeed, $\varphi(d^{1/(d+1)}) = D$, but for any $u > 0$ we have by the inequality between the arithmetic and geometric means

$$\frac{u^{-d} + u}{d+1} \geq \left(u^{-d}\frac{u}{d} \cdots \frac{u}{d}\right)^{\frac{1}{d+1}} = d^{-\frac{d}{d+1}} = \frac{D}{d+1}.$$

It follows that

$$\begin{aligned}\int_0^\infty \exp\left(\nu\left\{-\frac{1}{u^d} - u\right\}\right) du &< \int_0^D e^{-\nu D} du + \int_D^\infty e^{-\nu u} du \\ &= De^{-\nu D} + e^{-\nu D}\int_0^\infty e^{-\nu u} du = \left(D + \frac{1}{\nu}\right) e^{-\nu D}.\end{aligned}$$

As $\nu \to \infty$, this remains $O(e^{-\nu D})$, which is quite sufficient for our present purposes. □

Comparing (74), (75) and (77) we are led to

$$\left|\int_\gamma \left(-\frac{\zeta'(s)}{\zeta(s)}\right) \frac{x^s}{s^2} ds\right| \ll x \log^A T \sqrt{\nu} e^{-D\nu}, \tag{78}$$

with $D$ and the implied constant of $\ll$ depending explicitly on $d$. So substituting the definition $\nu = (a \log x)^{1/(d+1)}$ we finally get

$$\left|\int_\gamma \left(-\frac{\zeta'(s)}{\zeta(s)}\right) \frac{x^s}{s^2} ds\right| \ll x \log^A T \log^{\frac{1}{2d+2}} x e^{-D(a \log x)^{\frac{1}{d+1}}} \quad \text{with} \quad D = \frac{d^{\frac{1}{d+1}}(d+1)}{d}. \tag{79}$$

To finish the proof we choose $T = x^2$, say, and collect the error terms (72), (73) and (79) to get

$$\left| \int_\Gamma \left( -\frac{\zeta'(s)}{\zeta(s)} \right) \frac{x^s}{s^2} ds \right| \ll \sqrt{x} + x \log^{A+1} x e^{-D(a \log x)^{\frac{1}{d+1}}} = O\left( x e^{-(D-\epsilon)(a \log x)^{\frac{1}{d+1}}} \right), \tag{80}$$

with $\epsilon > 0$ arbitrary and $x > x_0(\epsilon)$ large enough.

Combining (69), (70) and (80) yields

$$T(x) = x + O\left( x e^{-(D-\epsilon)(a \log x)^{\frac{1}{d+1}}} \right). \tag{81}$$

To deduce Theorem 3, take now $b < a^{1/(d+1)} D$ arbitrary. Choosing $\epsilon > 0$ sufficiently small, (81) immediately implies

$$T(x) = x + O\left(\delta(x)\right) \qquad \left( \delta(x) := x \exp(-b \log^{\frac{1}{d+1}} x) \right). \tag{82}$$

Now an application of Lemma 7 leads to the assertion. □

To prove Theorem 2, we substitute $d = 1$, and find $D = 2$. Whence with all constants $b < 2\sqrt{a}$, the asymptotic formula (68) holds true. Theorem 1 is even simpler to obtain.

## §8. Application of positive polynomials to find zero-free regions of $\zeta(s)$

This is the point where positive trigonometric polynomials – and the quantities $U$ and $V$ of the Landau problems – have their role. In this section let

$$f(x) := a_0+a_1 \cos x+\cdots+a_k \cos kx \geq 0 \qquad 0 \leq a_j \ (j=1,\ldots,k) \ 0 < a_0 < a_1 \quad (83)$$

be any positive, positive definite even trigonometric polynomial with first coefficient $a_1$ larger than its constant term $a_0$.

**Proposition 10 (Landau).** *Let $f$ be any polynomial in* (83), *and let $b$ be any constant with*

$$b \geq \frac{f(0)}{a_1 - a_0} - 1\,.$$

*Then we have*

$$|\zeta(\sigma+it)| \geq c \log^{-b}|t| \qquad |t| \geq 2, \quad \sigma = 1 + \frac{1}{\log^{b+2}|t|}.$$

*Proof.* It is easy to see either directly from the definition and then the Euler product of $\zeta(s)$ or from integrating $\zeta'/\zeta$ that we have

$$\log(\zeta(s)) = \sum_p \sum_\ell \frac{1}{\ell} p^{-\ell s} = \sum_p \sum_\ell \frac{1}{\ell} p^{-\sigma \ell} e^{-it\ell \log p} \qquad (\sigma > 1).$$

Taking also the conjugate point $\overline{s}$ and applying $2\Re \log z = \log z + \log \overline{z} = \log |z|^2 = 2\log|z|$ we find

$$\log|\zeta(s)| = \Re \log \zeta(s) = \sum_p \sum_\ell \frac{1}{\ell} p^{-\ell\sigma} \cos(t\ell \log p). \tag{84}$$

Combining (84) at the points $\sigma + ijt$ with the coefficients $a_j$ we obtain

$$\log \prod_{j=0}^{k} |\zeta(\sigma+ijt)|^{a_j} = \sum_p \sum_\ell \frac{1}{\ell} p^{-\ell\sigma} f(\ell \log p) \geq 0, \tag{85}$$

that is,

$$\prod_{j=0}^{k} |\zeta(\sigma+ijt)|^{a_j} \geq 1. \tag{86}$$

For a lower estimate of $|\zeta(s)|$ we apply Proposition 1 (iv) and Proposition 3 (i) for the other terms to get

$$|\zeta(s)|^{a_1} \geq c\left(\frac{\sigma-1}{\sigma}\right)^{a_0}\left(\frac{1}{\log t}\right)^{a_2+\cdots+a_k}$$
$$\gg \left(\frac{1}{\log t}\right)^{(b+2)a_0+a_2+\cdots+a_k} \qquad \left(t \geq 2, \quad \sigma = 1+\frac{1}{\log^{b+2} t}\right).$$

Thus we get

$$\left(|\zeta(s)|\log^b t\right)^{a_1} \gg (\log t)^{ba_1-((b+2)a_0+a_2+\cdots+a_k)},$$

hence we need to choose $b$ so that the exponent on the right hand side be nonnegative. A calculation leads

$$b(a_1-a_0) \geq f(0)-(a_1-a_0),$$

which is equivalent to the condition assumed for $b$. This proves the Proposition. □

Now we can combine Proposition 4, Proposition 9 and 10 with Theorem 1 to conclude

**Corollary 1 (Landau).** *Let*

$$U := \inf\left\{\frac{f(0)}{a_1-a_0} \ : \ f \text{ satisfies } (83)\right\}.$$

*Then we have with any $K>0$ and $L<1/(U+1)$ the formula* (2).

**Remark 1.** Note the improvement of obtaining $1/(U+1)$ and not $1/(U+2)$ originally derived by Landau. This is due to the fine estimates of §4, where we have improved a root factor upon Landau's original derivation.

The above method yields an error term $\exp(\log^L x)$, still not optimal. Applying the even sharper estimates of $\zeta'/\zeta$ in Propositions 6 and 8 should further improve our end result. In fact, already de la Vallée Poussin achieved better, and an improvement upon that belongs to the second, more precise approach of Landau.

**Proposition 11 (Landau).** *Let $f$ be any polynomial in* (83), *and let $a$ be any constant with*

$$a < \frac{2(\sqrt{a_1}-\sqrt{a_0})^2}{f(0)-a_0}.$$

*Then we have*

$$|\zeta(\sigma+it)| \neq 0 \qquad \text{whenever} \qquad |t| \geq t_0(a), \quad \sigma \geq 1-\frac{a}{\log|t|}.$$

*Proof.* Now we combine the Dirichlet series representation of the logarithmic derivative of $\zeta(s)$ at the points $\sigma + ijt$ with coefficients $a_j$ and take real parts to get

$$\Re \sum_{j=0}^{k} -a_j \frac{\zeta'}{\zeta}(\sigma + ijt) = \sum_{j=0}^{k} a_j \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^\sigma} \Re e^{-ijt\log n} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^\sigma} \sum_{j=0}^{k} a_j \cos(jt\log n), \tag{87}$$

which implies

$$\Re \sum_{j=0}^{k} a_j \left( -\frac{\zeta'}{\zeta}(\sigma + ijt) \right) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^\sigma} f(t \log n) \geq 0. \tag{88}$$

In the other direction we apply Proposition 8 for all $j \geq 1$ with choosing the set of zeroes to be taken into account simply the empty set if $j \geq 2$, and the only one zero at $\rho = \beta + it$ lying closest to the critical line at height $t$ when $j = 1$. (We present the argument only for imaginary parts $t = \Im\rho$ where there is at least one root $\rho$, for otherwise there is nothing to prove.) Moreover, formula (45) gives a similar estimate $-\zeta'(\sigma)/\zeta(\sigma) \leq 1/(\sigma - 1) + O(1)$ for $1 < \sigma < 2$. These give

$$0 \leq \Re \sum_{j=0}^{k} a_j \left( -\frac{\zeta'}{\zeta}(\sigma + ijt) \right) \leq \frac{a_0}{\sigma - 1} - \frac{a_1}{\sigma - \beta} \frac{a_1 + a_2 + \ldots a_k}{2} \log t + O(1). \tag{89}$$

Put now

$$\kappa := \frac{a_0}{a_1} \in (0,1) \qquad \text{and} \qquad \lambda = \frac{a_1 + a_2 + \ldots a_k}{2a_1} + \eta \quad \text{with} \quad \eta > 0 \quad \text{arbitrary.}$$

We can melt the $O(1)$ term into $\eta \log t$ for $t > t_0$, whence

$$\frac{1}{\sigma - \beta} < \kappa \frac{1}{\sigma - 1} + \lambda \log t \qquad (t > t_0).$$

Put now with a parameter $\mu > 0$

$$\sigma = 1 + \frac{\mu}{\log t}.$$

It follows that

$$\frac{1}{1 - \beta + \frac{\mu}{\log t}} < \left( \frac{\kappa}{\mu} + \lambda \right) \log t \qquad (t > t_0),$$

$$1 - \beta + \frac{\mu}{\log t} > \frac{1}{\frac{\kappa}{\mu} + \lambda} \frac{1}{\log t} \qquad (t > t_0),$$

$$1 - \beta > \left( \frac{1}{\frac{\kappa}{\mu} + \lambda} - \mu \right) \frac{1}{\log t} = \frac{1 - \kappa - \lambda\mu}{\frac{\kappa}{\mu} + \lambda} \frac{1}{\log t} \qquad (t > t_0).$$

We wish to choose the parameter $\mu$ – that is, the value of $\sigma$ – so that this last coefficient of $1/\log t$ be the largest possible. Calculus gives

$$\mu := \frac{\sqrt{\kappa} - \kappa}{\lambda},$$

a positive quantity since $0 < \kappa < 1$. Substituting this choice of $\mu$ into the above expression, we find

$$\frac{1 - \kappa - \lambda\mu}{\frac{\kappa}{\mu} + \lambda} = \frac{1 - \sqrt{\kappa}}{\frac{\kappa\lambda}{\sqrt{\kappa}-\kappa} + \lambda} = \frac{(1 - \sqrt{\kappa})(\sqrt{\kappa} - \kappa)}{\kappa\lambda + (\sqrt{\kappa} - \kappa)\lambda} = \frac{(1 - \sqrt{\kappa})^2}{\lambda},$$

and our estimate becomes

$$1 - \beta > \frac{(1 - \sqrt{\kappa})^2}{\lambda} \frac{1}{\log t} \qquad (t > t_0).$$

That is,

$$1 - \beta > \frac{a}{\log t} \qquad \text{if} \quad a < \frac{\left(1 - \sqrt{\frac{a_0}{a_1}}\right)^2}{\frac{a_1 + a_2 + \ldots a_k}{2a_1}} = \frac{2(\sqrt{a_1} - \sqrt{a_0})^2}{a_1 + a_2 + \ldots a_k} \quad \text{and} \quad t > t_0(a).$$

The assertion is proved. □

As before, we now combine the zero-free domain obtained and the general derivation of the corresponding error term in Theorem 2 to obtain

**Corollary 2 (Landau).** *Let*

$$V := \inf\left\{ \frac{f(0) - a_0}{(\sqrt{a_1} - \sqrt{a_0})^2} \ : \ f \ \textit{satisfies} \ (83) \right\}.$$

*Then we have with $L = 1/2$ and any $K < 2\sqrt{2}/\sqrt{V}$ the formula* (2).

*Proof.* Combining the zero-free region of Proposition 11 and the logarithmic derivative estimate of Proposition 9 furnishes condition (67) with $d = 1$ and $a < 2/V$ arbitrary. Thus an application of Theorem 3 yields the assertion with $K := b < 2\sqrt{a} = 2\sqrt{2}/\sqrt{V}$. □

**Remark 2.** Such type of extremal estimates still play a vital role in analytic number theory. D. R. Heathbrown [14] uses similar combinations of trigonometric polynomials in his work on the Linnik constant; J. Pintz has asked recently the question of finding optimal polynomials for the expression

$$\inf\left\{ \frac{a_0 + f(0)/9}{(\sqrt{a_1} - \sqrt{a_0})^2} \right\},$$

etc.

## §9. The history of Landau's extremal problems

**9.1.** In the following we deal with the extremal quantities of Landau. Our discussion will essentially follow the paper [23] – for more details and various comments about contribution of colleagues on the development of our work in the subject see the original paper.

As attracting many eminent mathematicians, the determination, or estimation of $U$ and $V$ became a well-known problem independently of its number theoretic applications. To give a historical account of results thus far, let us introduce a more systematic notation. We put for any $a \in \mathbb{R}$

$$\mathcal{F}(a) := \Bigg\{ f \in C(\mathbb{T}) : f(x) = 1 + a\cos x + \sum_{k=2}^{\infty} a_k \cos kx \geq 0 \quad (\forall x), \qquad a_k \geq 0 \quad (k \in \mathbb{N}) \Bigg\} \tag{90}$$

and denote

$$\begin{aligned} &\mathcal{F}_n(a) := \mathcal{F}(a) \cap \mathcal{T}_n, \quad \mathcal{F}^*(a) := \mathcal{F}(a) \cap \mathcal{T}, \quad \mathcal{F} = \bigcup_{a>1} \mathcal{F}(a), \\ &\mathcal{F}_n := \bigcup_{a>1} \mathcal{F}_n(a) = \mathcal{F} \cap \mathcal{T}_n, \quad \mathcal{F}^* := \bigcup_{a>1} \mathcal{F}^*(a) = \mathcal{F} \cap \mathcal{T} = \bigcup_{n=1}^{\infty} \mathcal{F}_n. \end{aligned} \tag{91}$$

One can define

$$\begin{aligned} \alpha(a) &:= \inf \big\{ f(0) : f \in \mathcal{F}(a) \big\}, \\ \alpha^*(a) &:= \inf \big\{ f(0) : f \in \mathcal{F}^*(a) \big\}, \\ \alpha_n(a) &:= \inf \big\{ f(0) : f \in \mathcal{F}_n(a) \big\}. \end{aligned} \tag{92}$$

Note that the definitions (92) can be used whenever $\mathcal{F}(a) \neq \emptyset$, $\mathcal{F}^*(a) \neq \emptyset$ or $\mathcal{F}_n(a) \neq \emptyset$, resp. It is easy to see that $\alpha(a) = \alpha^*(a)$ for all $a \in \mathcal{D}(\alpha)$ except possibly for the point $A$ at the left end of the domain of $\alpha$ where $\mathcal{F}^*(A)$ may be empty. It is also easy to see that $[1,2) \subset \mathcal{D}(\alpha)$, $\alpha(a)$ is continuous in $[1,2)$, and that $\mathcal{F}(a) = \emptyset$ for $a \geq 2$; moreover, $\alpha(a) \to +\infty$ as $a \to 2-0$. Finally the infimum in the definition of $\alpha(a)$ is actually a minimum,

$$\begin{aligned} \alpha(a) &= \min \big\{ f(0) : f \in \mathcal{F}(a) \big\}, \\ \alpha_n(a) &= \min \big\{ f(0) : f \in \mathcal{F}_n(a) \big\}. \end{aligned} \tag{93}$$

These observations can be found in [22] in a more general setting. However, we have to note that most of the facts mentioned here appeared first in [2] where $\chi(a) = \alpha(a) - 1$ and $\chi_n(a) = \alpha_n(a) - 1$ are defined and analyzed (for $a \geq 1$). This analysis is continued (for $a \geq 0$) in [1].

Thus we can omit $\alpha^*$ and $\mathcal{F}^*$ from now on using only $\alpha$ and $\mathcal{F}$ in place of the original, equivalent usage of Landau. With these cleared we can also put

$$\begin{aligned} U &:= \min_{a>1} \frac{\alpha(a)}{a-1}, & U_n &:= \min_{a>1} \frac{\alpha_n(a)}{a-1}, \\ V &:= \min_{a>1} \frac{\alpha(a)-1}{(\sqrt{a}-1)^2}, & V_n &:= \min_{a>1} \frac{\alpha_n(a)-1}{(\sqrt{a}-1)^2}, \end{aligned} \tag{94}$$

where the use of min in place of inf is justified later.

Plainly for all $a$ we have $\alpha_n(a) \searrow \alpha(a)$ $(n \to \infty)$ and $U_n \searrow U$, $V_n \searrow V$ $(n \to \infty)$. (Here $\searrow$ means monotonically nonincreasing convergence.) Below is a list of values already determined.

| | | |
|---|---|---|
| $U_2 = 7$ | Landau [18] & Chakalov, [4, 5] | |
| $U_3 = U_4 = U_5 = 6$ | Landau [19, 20] & Chakalov, [4, 5] | |
| $U_6 = 5.92983\ldots$ | Chakalov, [4, 5] | |
| $U_7 = U_8 = U_9 = 5.90529\ldots$ | Chakalov, [4, 5] | (95) |
| $V_2 = 53.1390719\ldots$ | French, [13] | |
| $V_3 = 36.9199911\ldots$ | Arestov, [1] | |
| $V_4 = V_5 = V_6 = 34.8992258\ldots$ | Arestov, [1]. | |

Estimates were also deduced for many of the extremal quantities. It follows a list of records to date in estimating these values.

| | | |
|---|---|---|
| $U < 5.90529\ldots$ | Chakalov, [4, 5] | |
| $U > 5.8726$ | Arestov–Kondrat'ev, [2] | |
| $V < 34.5035864\ldots$ | Arestov–Kondrat'ev, [2] | (96) |
| $V > 34.468305\ldots$ | Arestov–Kondrat'ev, [2] | |
| $V_8 < 34.54461566$ | Kondrat'ev, [16]. | |

For historical completeness let us mention a few other results, already improved upon.

| | |
|---|---|
| $V < 35.074$ | Westphal, [33] |
| $V_2 \le 53.15$ | Stechkin, [29] |
| $U_{11} > 5.792$ | Chakalov, [4, 5] |
| $V_3 \le 37.04$ | Landau, [18] |
| $V_3 < 36.97$ | Stechkin, [29] |
| $V_4 < 35.03264$ | Rosser and Schönfeld, [26] |

| | |
|---|---|
| $V > 21.64$ | French, [13] referring to an unpublished result of Schoenfeld & V. J. Le Veque |
| $V > 32.49$ | Stechkin, [29] |
| $U > 5.8642$ | B. L. van der Waerden, [31] |
| $V > 32.5136$ | French, [13] |
| $36.96 > V_3 > 36.59$ | Bateman (unpublished, quoted in [13]) and [29], resp. |
| $34.91 > V_4 > 34.35$ | Stechkin, [29] |
| $V_{5,6,7} > 33.373$ | French, [13] |
| $V_8 > 33.313$ | French, [13] |
| $V_9 > 33.1766$ | French, [13] |

$$V_n \geq \left(8 - \frac{3\pi - 7}{2\cos\left(\frac{\pi}{n+2}\right) - 1}\right) \cdot \frac{\sqrt{2\cos\left(\frac{\pi}{n+2}\right) + 1}}{\sqrt{2\cos\left(\frac{\pi}{n+2}\right) - 1}} \quad (n \in \mathbb{N}) \quad \text{Stechkin, [29]}$$

| | |
|---|---|
| $34.8993 > V_4$ | D. Hollenbeck (unpublished, referred to in [27]) |
| $V > 33.58$ | Reztsov, [24] |
| $U > 5.8656$ | Révész, unpublished |
| $33.54 < V < 34.677$ | Révész, unpublished. |

**9.2.** In the many investigations of Landau's extremal problems, a number of new relatives were introduced. In his quite elegant and sharp lower estimation for $U$, van der Waerden [31] used the construction of a measure

$$d\kappa(x) \sim b_0 + 2\sum_{k=1}^{\infty} b_k \cos kx \geq 0 \tag{97}$$

with the properties

$$\kappa \geq 0, \quad b_0 + b_1 \leq 2, \quad b_k \leq 1 \quad (k \in \mathbb{N}_2), \tag{98}$$

where $\mathbb{N}_2 := \mathbb{N} \cap [2, \infty]$. Actually van der Waerden sought minimal $b_1$ ($b_1 < 0$ with maximal absolute value) and could prove that $U \geq 1 - b_1$. Formulating this as an extremal problem, van der Waerden treated

$$\Omega := \sup\left\{1 - b_1 : \exists\kappa \in BM(\mathbb{T}), \quad \kappa \geq 0 \quad \text{with (97)–(98)}\right\}. \tag{99}$$

Finding a measure with (97)–(98) and with $b_1 = -4.8642\ldots$, van der Waerden showed actually

$$U \geq \Omega \geq 5.8642\ldots . \tag{100}$$

S. B. Stechkin [29] used a different method aiming mainly the estimation of $V$. In the course of proof he defined an intermediate quantity between $U$ and $V$ when introducing

$$\begin{aligned} W &:= \inf\left\{\frac{f(0)-1}{a_1-1} : f \in \mathcal{F}(a_1) \quad \text{with } a_1 > 1\right\} = \min_{a>1} \frac{\alpha(a)-1}{a-1}, \\ W_n &:= \inf\left\{\frac{f(0)-1}{a_1-1} : f \in \mathcal{F}_n(a_1) \quad \text{with } a_1 > 1\right\} = \min_{a>1} \frac{\alpha_n(a)-1}{a-1}. \end{aligned} \tag{101}$$

As in case of $U$ and $V$, we again have $W_n \searrow W$, and the determination of $W$ and $W_n$ is a problem of a similar sort.

Stechkin himself could estimate $W$ as follows.

$$\begin{aligned} &W_2 = \frac{1}{2}\left(5+\sqrt{17}\right) = 4.56\ldots, \\ &W_4 \leq W_3 \leq \frac{1}{2}\left(5+\sqrt{13}\right) = 4.30\ldots, \\ &W \geq 4.159. \end{aligned} \tag{102}$$

**9.3.** We also introduce some more extremal quantities. Denote by $\lambda$ and $\delta_z$ ($z \in \mathbb{T}$) the normalized measures (the Lebesgue and (essentially) the Dirac measures at $z \in \mathbb{T}$)

$$\begin{aligned} d\lambda(x) &\sim 1, \\ d\delta_z(x) &\sim 1 + 2\sum_{b=1}^{\infty}\big(\cos kz\cos kx + \sin kz \sin kx\big), \\ \delta &:= \delta_0. \end{aligned} \tag{103}$$

We consider the measure sets

$$\begin{aligned} \mathcal{M}(0) :=& \Bigg\{\tau \in BM(\mathbb{T}) : d\tau(x) \sim \sum_{k=1}^{\infty} t_k \cos kx, \\ &\qquad t_1 \in \mathbb{R}, \quad t_k \leq 0 \quad (k \in \mathbb{N}_2)\Bigg\}, \\ \mathcal{M}_n(0) :=& \Bigg\{\tau \in BM(\mathbb{T}) : d\tau(x) \sim \sum_{k=1}^{\infty} t_k \cos kx, \\ &\qquad t_1 \in \mathbb{R}, \quad t_k \leq 0 \quad (2 \leq k \leq n)\Bigg\}, \end{aligned}$$

$$
\begin{aligned}
\mathcal{M}(a) :=& \Bigg\{ \tau \in BM(\mathbb{T}) : d\tau(x) \sim b\left(1 - \frac{2}{a}\cos x\right) + \sum_{k=2}^{\infty} t_k \cos kx, \\
& \qquad b \in \mathbb{R}, \quad t_k \le 0 \quad (k \in \mathbb{N}_2) \Bigg\} \qquad (104)\\
=& \left\{ \tau \in BM(\mathbb{T}) : \tau = -\frac{a}{2} t_1(\tau_0) \cdot \lambda + \tau_0, \quad \tau_0 \in \mathcal{M}(0) \right\}, \\
& \qquad\qquad \Big( t_1(\tau_0) := \langle \tau_0, 2\cos x \rangle \Big), \\
\mathcal{M}_n(a) :=& \Bigg\{ \tau \in BM(\mathbb{T}) : d\tau(x) \sim b\left(1 - \frac{2}{a}\cos x\right) + \sum_{k=2}^{\infty} t_k \cos kx, \\
& \qquad b \in \mathbb{R}, \quad t_k \le 0 \quad (2 \le k \le n) \Bigg\} \\
=& \left\{ \tau \in BM(\mathbb{T}) : \tau = -\frac{a}{2} t_1(\tau_0) \cdot \lambda + \tau_0, \quad \tau_0 \in \mathcal{M}_n(0) \right\}, \\
& \qquad\qquad \Big( t_1(\tau_0) := \langle \tau_0, 2\cos x \rangle \Big),
\end{aligned}
$$

where the last two definitions are valid for any $a \in \mathbb{R}$ and $a \neq 0$. Also we put for arbitrary $y \in \mathbb{R}$

$$
\begin{aligned}
\mathcal{N}(y) :=& \Bigg\{ \nu \in BM(\mathbb{T}) : \nu \ge 0, \quad d\nu(x) \sim 1 + \sum_{k=1}^{\infty} y_k \cos kx, \\
& \qquad y_1 \in \mathbb{R}, \quad y_k \le y \quad (k \in \mathbb{N}_2) \Bigg\}, \\
\mathcal{N}_n(y) :=& \Bigg\{ \nu \in BM(\mathbb{T}) : \nu \ge 0, \quad d\nu(x) \sim 1 + \sum_{k=1}^{\infty} y_k \cos kx, \\
& \qquad y_1 \in \mathbb{R}, \quad y_k \le y \quad (2 \le k \le n) \Bigg\}.
\end{aligned}
\qquad (105)
$$

Finally let us introduce for all $b \in (-2, 2)$ the square-integrable function set

$$
\mathcal{G}(b) := \left\{ g \in L^2(\mathbb{T}) : g \ge 0, \quad g(x) \sim 1 + b\cos x + \sum_{k=2}^{\infty} b_k \cos kx \right\}. \qquad (106)
$$

To these sets we define the following extremal quantities.

$$
\begin{aligned}
\omega(a) &:= \sup\big\{t : \exists \tau \in \mathcal{M}(a), \quad \tau + \delta \geq t \cdot \lambda\big\}, \\
\omega_n(a) &:= \sup\big\{t : \exists \tau \in \mathcal{M}_n(a), \quad \tau + \delta \geq t \cdot \lambda\big\}, \\
\beta(y) &:= \sup\big\{ - y_1 : \exists \nu \in \mathcal{N}(y), \quad y_1 = \langle \nu, 2\cos x\rangle\big\}, \\
\beta_n(y) &:= \sup\big\{ - y_1 : \exists \nu \in \mathcal{N}_n(y), \quad y_1 = \langle \nu, 2\cos x\rangle\big\}, \\
\vartheta(y) &:= \sup\big\{y_1 : \exists \nu \in \mathcal{N}(y), \quad y_1 = \langle \nu, 2\cos x\rangle\big\}, \\
\vartheta_n(y) &:= \sup\big\{ - y_1 : \exists \nu \in \mathcal{N}_n(y), \quad y_1 = \langle \nu, 2\cos x\rangle\big\}, \\
\gamma(b) &:= \inf\big\{||g||_2 : g \in \mathcal{G}(b)\big\}.
\end{aligned}
\tag{107}
$$

## §10. Preliminaries for the detailed analysis of $U$ and $V$

**10.1.** As we will extensively use sets of Borel measures and extremal quantities defined on these sets, we summarize a few facts of the structure of $BM(\mathbb{T})$ at the outset.

Let us recall that $BM(\mathbb{T}) = C(\mathbb{T})^*$, the topological dual of the Banach space $C(\mathbb{T})$ with the norm of the total variation norm

$$||\mu||_{BM(\mathbb{T})} = \int\limits_{\mathbb{T}} |d\mu|. \tag{108}$$

We know that $C(\mathbb{T})$ is not reflexive, and $BM(\mathbb{T})^* \supsetneqq C(\mathbb{T})$. Hence the weak, and the weak $*$ topology of $BM(\mathbb{T})$ are different, the weak topology being the weakest topology so that all functionals from $BM(\mathbb{T})^*$ be continuous linear functionals on $BM(\mathbb{T})$, while the weak $*$ topology is the weakest topology so that the functionals belonging to $C(\mathbb{T})$ be continuous on $BM(\mathbb{T})$. Thus the weak $*$ topology is even weaker than the weak one.

In a topological vector space convex and closed sets remain convex and closed when considering the weak topology in place of the original topology. However, in dual spaces like $BM(\mathbb{T})$ closedness is not necessarily saved when considering the weak $*$ topology instead of the weak topology. On the other hand we have the Banach–Alaoglu Theorem ([10], 4.10.3. Theorem, p. 205) stating that all the closed balls in the dual space $BM(\mathbb{T})$ are weak $*$ compact.

Our application of these structural facts will have the following pattern. Usually we define a set of measures in $BM(\mathbb{T})$ and wish to extremalize some quantity on that set. Using the definition, we can pass on to a decreasing sequence of closed, bounded and convex sets $F_n \subset BM(\mathbb{T})$, and to show that there exists an extremal measure, we are entitled to show that $F := \bigcap_{n=1}^{\infty} F_n \neq \emptyset$. This is a Cantor type property, and can be guaranteed for decreasing and nonempty sequences of compact sets. Now usually $F_n \subset BM(\mathbb{T})$ will be convex and closed, but not compact. To save the idea, we pass on to the weak $*$ topology. First, the nonempty sets $F_n \subset BM(\mathbb{T})$ remain convex in any topology. They will be bounded in the norm of $BM(\mathbb{T})$ usually because for nonnegative measures

$$||\mu||_{BM(\mathbb{T})} = \int\limits_{\mathbb{T}} |d\mu| = \int\limits_{\mathbb{T}} d\mu = 2\pi\langle 1, \mu\rangle \quad (\mu \in BM(\mathbb{T}), \quad \mu \geq 0). \tag{109}$$

Hence $F_n$ will be conditionally compact in the weak $*$ topology according to the Banach–Alaoglu Theorem. To show that $F_n$ are weak $*$ compact, the key point is to show that $F_n$ are weak $*$ closed, too.

It is obvious that any closed and convex sets $F_n$ can be represented as the intersection of a set of closed halfspaces defined by continuous linear functionals *from the bidual space.* However, such level sets of linear functionals can be proved to be even weak $*$ closed only if the functionals themselves are weak $*$ continuous, i.e. if the functionals belongs to $C(\mathbb{T})$. Thus we will look for a representation of $F_n$ as an intersection of level sets of the type

$$X(f,c) := \big\{\mu \in BM(\mathbb{T}) : \langle f, \mu \rangle \leq c\big\} \tag{110}$$

with $f \in C(\mathbb{T})$. Having such a representation, we can claim $F_n$ to be even weak $*$ closed, hence we get that $F_n$ is not only conditionally compact, but it is also compact in the weak $*$ topology. Finally, we can refer to the Cantor type property that the intersection of the compact, decreasing and nonempty sets $F_n$ must be nonempty. To formalize this argument, we can state the following.

**Lemma 13.** *Suppose that $F_n$ $(n \in \mathbb{N})$ is a sequence of subsets of $BM(\mathbb{T})$ with the following properties.*

i) *$F_n \neq \emptyset$ $(n \in \mathbb{N})$.*

ii) *$F_{n+1} \subset F_n$ $(n \in \mathbb{N})$.*

iii) *$F_n$ is bounded in the total variation norm of $BM(\mathbb{T})$ (perhaps for $n > n_0$).*

iv) *$F_n$ can be represented as the intersection of a number of closed halfspaces of the form* (110) *with the generating functionals belonging to $C(\mathbb{T})$.*

*Then the intersection*

$$F := \bigcap_{n=1}^{\infty} F_n \tag{111}$$

*is a norm-bounded, closed, convex, weak $*$-compact and nonempty subset of $BM(\mathbb{T})$.*

Consider the sets

$$\begin{aligned} BM(\mathbb{T})_C :=& \big\{\mu \in BM(\mathbb{T}) : \mu \\ &\qquad \text{is even (i.e. } \mu(H) = \mu(-H) \quad (\forall H \subset \mathbb{T}, \text{ measurable}))\big\} \\ =& \big\{\mu \in BM(\mathbb{T}) : \langle \sin kx, \mu \rangle = 0 \quad (k \in \mathbb{N})\big\} \\ =& \big\{\mu \in BM(\mathbb{T}) : \langle f, \mu \rangle = 0 \quad \forall f \in C(\mathbb{T}), \quad f(x) \equiv -f(-x)(x \in \mathbb{T})\big\}, \end{aligned} \tag{112}$$

$$\begin{aligned} BM(\mathbb{T})_S :=& \big\{\mu \in BM(\mathbb{T}) : \mu \\ &\qquad \text{is odd } (\mu(H) = -\mu(-H) \quad (\forall H \subset \mathbb{T}, \text{ measurable}))\big\} \\ =& \big\{\mu \in BM(\mathbb{T}) : \langle \cos kx, \mu \rangle = 0 \quad (k \in \mathbb{N})\big\} \\ =& \big\{\mu \in BM(\mathbb{T}) : \langle f, \mu \rangle = 0 \quad \forall f \in C(\mathbb{T}), \quad f(x) \equiv -f(-x)(x \in \mathbb{T})\big\}, \end{aligned} \tag{113}$$

and the set

$$BM(\mathbb{T})_P := \big\{\mu \in BM(\mathbb{T}) : \mu \text{ is nonnegative (i.e. } \mu(H) \geq 0 \ \ (\forall H \subset \mathbb{T}, \text{ measurable}))\big\} = \big\{\mu \in BM(\mathbb{T}) : \langle f, \mu\rangle \geq 0 \quad \forall f \in C(\mathbb{T}), \quad f \geq 0\big\}. \tag{114}$$

Here we use the sets of even, odd and nonnegative functions

$$\begin{aligned} E &:= \Big\{ f \in C(\mathbb{T}) : f(x) \equiv f(-x) \quad (x \in \mathbb{T}) \Big\}, \\ O &:= \Big\{ f \in C(\mathbb{T}) : f(x) \equiv -f(-x) \quad (x \in \mathbb{T}) \Big\}, \\ P &:= \Big\{ f \in C(\mathbb{T}) : f \geq 0 \Big\}, \end{aligned} \tag{115}$$

to establish a representation of the type iv) for the sets (112)–(114). Namely,

$$BM(\mathbb{T})_C = \bigcap_{f \in O} (X(f,0) \cap X(-f,0)), \tag{116}$$

$$BM(\mathbb{T})_S = \bigcap_{f \in E} (X(f,0) \cap X(-f,0)), \tag{117}$$

$$BM(\mathbb{T})_P = \bigcap_{f \in P} (X(-f,0). \tag{118}$$

Thus in the following we can use property iv) for the sets (112), (113) and (114).

**10.2.** Let $0 < a < b$, and $k : [a,b] \to \mathbb{R}$ be any continuous, strictly increasing and concave function on the interval.

We define the "tangential function to $k$" and the "extremal tangential curve" to $k$ as follows.

**Definition 1.** For $t \in \mathbb{R}$ let us consider the points $(x,t)$, $(0,t)$, $(x,k(x))$ in this order for all $x \in [a,b]$ and denote $\varphi(t,x)$ the angle (measured from the positive $x$ direction to the counterclockwise sense) of the chord drawn from $(0,t)$ to $(x,k(x))$. As $0 < a \leq x \leq b$, and the vector $((0,t),\ (x,t))$ is horizontal, we plainly have $-\frac{\pi}{2} < \varphi(t,x) < \frac{\pi}{2}$.

We introduce the "extremal tangential curve"

$$\Gamma := \Gamma_k := \Big\{(t_0,x_0) \in \mathbb{R}^2 : \varphi(t_0,x_0) = \max_{a \leq x \leq b} \varphi(t_0,x)\Big\}; \tag{119}$$

we also introduce the "tangential function to $k$"

$$f(t) := f_k(t) := \max_{a \leq x \leq b} \frac{k(x)-t}{x} = \max_{x \in [a,b]} \tan\varphi(t,x) = \tan\varphi(t,x^*) \qquad \big((t,x^*) \in \Gamma_k\big). \tag{120}$$

Plainly we may also consider

$$\varphi(t) := \arctan f(t) = \max_{a \le x \le b} \varphi(t,x) = \varphi(t,x^*) \quad ((t,x^*) \in \Gamma_k). \tag{121}$$

Geometrically $\varphi(t)$ is the oriented angle, $f(t)$ is the slope of the tangential straight line drawn from the point $(0,t)$ to the curve $\{(x,k(x)) : a \le x \le b\}$.

Note that $f_k(t)$ is just the well-known Legendre transform of the function $k$; as properties of the Legendre transform are well-known, see e.g. [25], in the following assertions we will omit the proofs.

**Lemma 14.** i) *The function $f(t) : \mathbb{R} \to \mathbb{R}$ is continuous and strictly decreasing.*

ii) *The curve $\Gamma$ is "oriented positively" in the sense that for any two points $(t',x')$, $(t'',x'') \in \Gamma \quad t' < t''$ entails $x' \le x''$.*

iii) *The point set $I(t) := \{x : (t,x) \in \Gamma\}$ is a convex closed set $\subset [a,b]$.*

Now let us define

$$\begin{aligned}
\underline{x}(t) &:= \min\big\{x : (t,x) \in \Gamma\big\} = \min I(t);\\
\overline{x}(t) &:= \max\big\{x : (t,x) \in \Gamma\big\} = \max I(t);\\
x(t) &:= \underline{x}(t) \quad \text{whenever} \quad \underline{x}(t) = \overline{x}(t);\\
T(x) &:= \big\{t \in \mathbb{R} : (t,x) \in \Gamma\big\};\\
\overline{t}(x) &:= \max\big\{t : (t,x) \in \Gamma\big\} = \max T(x);\\
\underline{t}(x) &:= \min\big\{t : (t,x) \in \Gamma\big\} = \min T(x);\\
t(x) &:= \underline{t}(x) \quad \text{whenever} \quad \underline{t}(x) = \overline{t}(x).
\end{aligned} \tag{122}$$

The existence and nature of $T$, $\overline{t}$, $\underline{t}$ are similar to $I$, $\overline{x}$, $\underline{x}$, by the very same Lemma 14 ii) and iii).

**Lemma 15.** i) *The concave function $k$ is differentiable iff $t(x)$ exists for all $a < x < b$. Moreover, for any $x \in (a,b)$ we have $\overline{t}(x) = \underline{t}(x)$ iff $k'(x-0) = k'(x+0)$.*

ii) *We always have*

$$f'(t+0) = \frac{-1}{\overline{x}(t)}, \quad f'(x-0) = \frac{-1}{\underline{x}(t)}.$$

**Corollary 3.** *If $f$ is the tangential function defined to $k$, then $f$ is always a continuous, strictly decreasing convex function. Moreover, $f$ is differentiable (and then also continuously) iff $k$ is strictly concave. Conversely, $f$ is strictly convex iff $k$ is differentiable iff $k \in C^1[a,b]$.*

## §11. Analysis of the extremal quantities

**11.1.** First of all let us record some basic properties of the functions defined in § 9. As for the domain of definition (where the corresponding definition yields a finite value) we use the notation $\mathcal{D}$; similarly, the range of a function is denoted by $\mathcal{R}$.

**Proposition 12.** i) *$\mathcal{D}(\alpha) = (A, 2)$ or $[A, 2)$, where $-\sqrt{3} \le A \le -\sqrt{2}$.*

ii) *$\mathcal{D}(\alpha_n) = [A_n, B_n]$, where $A_n \le -\sqrt{2}$ $(n \ge 2)$ and $B_n = 2\cos\frac{\pi}{n+2}$.*

iii) *$\mathcal{D}(\gamma) = (-2, 2)$.*

*Proof.* i) See Proposition 4.1 of [22].

ii) The first estimate follows from the example (4.2) of [22], and the second statement is a consequence of a theorem of Fejér [11] and Szász [28] who determined the corresponding extremal polynomials. See also [29], Lemma 1.

iii) Similar to i) but essentially trivial. □

**Proposition 13.** i) *In the definition* (92) *of $\alpha(a)$ the infimum is actually a minimum, i.e.*

$$\alpha(a) = \min\left\{f(0) : f \in \mathcal{F}(a)\right\}$$

*for all $a \in \mathcal{D}(\alpha)$.*

ii) *If $A \in \mathcal{D}(\alpha)$, then $\lim\limits_{a\to A+} \alpha(a) = \alpha(A)$, and if $A \notin \mathcal{D}(\alpha)$, then $\lim\limits_{a\to A+} \alpha(a) = \infty$.*

iii) *$\lim\limits_{a\to 2-} \alpha(a) = \infty$.*

iv) *$\alpha(a)$ is a convex function on $\mathcal{D}(\alpha)$.*

v) *$\alpha_n(a)$ is a convex function on $\mathcal{D}(\alpha_n)$.*

*Proof.* These can be found in [22], Propositions 4.2 and 4.3 or, in a somewhat more general form, in [21], 2.3, 2.5 and 2.6 Propositions. Note that i) and v) appeared already in Theorem 1 3) of [2], while iii) was proved first in [24], see also the comments to Proposition 16 iii). □

Our knowledge about the actual function values of $\alpha(a)$ is summarized in the next three propositions.

**Proposition 14.** i) *$\alpha(a) = 1 + a$ for $-1 \le a \le 1$.*

ii) *$\alpha(a) = 2a$ for $1 \le a \le 4/3$.*

iii) *$\alpha(a) = 0$ for $-4/3 \le a \le -1$.*

**Proposition 15.** *$\alpha(a) > 0$ for $a < -4/3$, $a \in \mathcal{D}(\alpha)$.*

**Proposition 16.** *In the range $4/3 < a < 2$ we have the following lower estimates for the function $\alpha(a)$.*

i) $\alpha(a) > \delta(a) := 2a$,

ii) $\alpha(a) \geq \varphi(a) := 8a - 3\pi$,

iii) $\alpha(a) \geq \tau(a) := \sqrt{\frac{2+a}{2-a}}$.

*Proofs.* For a proof of the claims in Proposition 14, see e.g. Proposition 4.4 i), ii) and iv) in [22]. Note that i) is (0.17) of [1] and ii) is covered by (0.15) of [1] or Theorem 1 1) in [2].

For Proposition 15 see Proposition 4.4 v) in [22].

Lastly, consider Proposition 16. First, i) can be found in Proposition 4.4 iii) of [22]. ($\alpha(a) \geq \delta(a)$ is trivial from Proposition 14 ii) and Proposition 13 iii).)

Proposition 15 ii) is an estimate of Stechkin, see Lemma 3 in [29].

Finally, the nontrivial estimate of Proposition 16 iii) is proved e.g. in [22], Theorem 5.1. A very similar proof of a very similar, but somewhat more elaborated (and thus slightly better) nonlinear estimate was given first in [24]. Actually, the key lemma to the result was attributed to Yudin (oral communication) in [24], while in [22] an independent and different proof was given which precisely characterizes also the extremal cases of the lemma. □

**11.2.** For the functions introduced in (107) their use and relevance to the problems studied can be best seen from the relation between $\alpha$ and $\omega$. Let $f \in \mathcal{F}(a)$ be arbitrary and take any $\tau \in \mathcal{M}(a)$ satisfying $\tau + \delta \geq t \cdot \lambda$ with some $t \geq 0$. (Such $t$ and $\tau$ must exist since the zero measure, $\mathbf{0} \in \mathcal{M}(a)$.) We have from the nonnegativity of $f$ and $a_k$ ($k \in \mathbb{N}_2$), the nonpositivity of $t_k$ ($k \in \mathbb{N}_2$) and from $\delta \geq t \cdot \lambda - \tau$, that

$$\begin{aligned} f(0) &= \langle f, \delta \rangle \geq \langle f, t \cdot \lambda - \tau \rangle \\ &= t - \langle f, \tau \rangle = t - \left\{ b + \frac{1}{2} \cdot a \cdot b \cdot \left( \frac{-2}{a} \right) + \sum_{k=2}^{\infty} a_k t_k \right\} = t - \sum_{k=2}^{\infty} a_k t_k \geq t. \end{aligned} \tag{123}$$

Taking supremum over all $\tau$ and $t$ on the right, and then infimum at the left-hand side, we obtain the inequality

$$\alpha(a) \geq \omega(a). \tag{124}$$

That estimate was essentially at the heart of van der Waerden's estimate, as we shall see later. This estimate is not only close numerically, but actually it is theoretically exact.

**Theorem 4.** (Duality)

i) $\mathcal{D}(\alpha) = \mathcal{D}(\omega)$, *and the* sup *in the first definition of* (107) *is actually a maximum.*

ii) *For all* $a \in \mathcal{D}(\alpha) \quad \alpha(a) = \omega(a)$.

iii) *For all* $n \in \mathbb{N} \quad \mathcal{D}(\alpha_n) = \mathcal{D}(\omega_n)$, *and the* sup *in the second definition of* (107) *is actually a maximum.*

iv) *For all* $n \in \mathbb{N}$ *and* $a \in \mathcal{D}(\alpha_n) \quad \alpha_n(a) = \omega_n(a)$.

*Proof.* The easy part is $\alpha(a) \geq \omega(a)$ and its relatives $\alpha_n(a) \geq \omega_n(a)$, as shown above. That also entails $\mathcal{D}(\alpha) \subset \mathcal{D}(\omega)$, $\mathcal{D}(\alpha_n) \subset \mathcal{D}(\omega_n)$. The converse is nontrivial, and the proof applies functional analysis. For the whole argument we refer to [21], especially 3.4 Theorem and 3.5 Proposition. Note that here the index sets $M$ and $L$ of [21] are $\mathbb{N}_2$ and $\emptyset$ or $[2, n]$ and $\emptyset$, and thus also 2.6 Proposition of [21] applies. That covers the border cases $a = A$ and $a = A_n$ or $B_n$, not included in the even more general setting of 3.4 Theorem of [21]. The existence of extremal measures $\omega$ and $\omega_n$ follows from the argument as pointed out in section 3.6 of [21]. $\square$

**11.3.** With the above duality theorem at hand, let us also define the functions

$$
\begin{aligned}
U(a) &:= \frac{\alpha(a)}{a-1} = \frac{\omega(a)}{a-1}, \\
V(a) &:= \frac{\alpha(a)-1}{(\sqrt{a}-1)^2} = \frac{\omega(a)-1}{(\sqrt{a}-1)^2}, \\
W(a) &:= \frac{\alpha(a)-1}{a-1} = \frac{\omega(a)-1}{a-1} \qquad (a \in (1,2)),
\end{aligned}
\tag{125}
$$

and for any $n \in \mathbb{N}$ their finite degree counterparts

$$
\begin{aligned}
U_n(a) &:= \frac{\alpha_n(a)}{a-1} = \frac{\omega_n(a)}{a-1}, \\
V_n(a) &:= \frac{\alpha_n(a)-1}{(\sqrt{a}-1)^2} = \frac{\omega_n(a)-1}{(\sqrt{a}-1)^2}, \\
W_n(a) &:= \frac{\alpha_n(a)-1}{a-1} = \frac{\omega_n(a)-1}{a-1} \qquad \big(a \in (1, B_n]\big).
\end{aligned}
\tag{126}
$$

Since the functions (125)–(126) are the products of one of the positive convex functions $\alpha(a)$, $\alpha(a)-1$, $\alpha_n(a)$, or $\alpha_n(a)-1$ and one of the strictly convex and positive functions $\frac{1}{a-1}$ or $\frac{1}{(\sqrt{a}-1)^2}$, all functions are positive and strictly convex. Note also that all the six functions tend to $+\infty$ as $a \to 1+0$ as the denominators tend to

$+0$ and the numerators are finite and positive. Similarly, as Proposition 16 iii) entails, $\alpha(a) \to +\infty$ $(a \to 2-0)$, and that implies $U(a) \to +\infty$, $V(a) \to +\infty$ and $W(a) \to +\infty$ $(a \to 2-0)$. Hence we see that the functions (125)–(126) all have minimum points where the extremal quantities (94) and (101) are attained, and also that these points are unique due to strict convexity. Thus we have

**Proposition 17.** i) *All the functions* (125)–(126) *are strictly positive and strictly convex in their domain of definition.*

ii) *All the functions* (125)–(126) *have limit* $+\infty$ *at* $1+0$.

iii) *The functions* (125) *have limit* $+\infty$ *at* $2-0$ *while the functions* (126) *are continuous and finite at* $B_n$.

iv) *The functions* (125)–(126) *have unique minimum points* $a_U$, $a_V$, $a_W$ *and* $a_{U,n}$, $a_{V,n}$, $a_{W,n}$, *respectively, where we have*

$$\begin{aligned} U &= U(a_U), & V &= V(a_V), & W &= W(a_W), \\ U_n &= U_n(a_{U,n}), & V &= V(a_{V,n}), & W &= W(a_{W,n}). \end{aligned}$$

**11.4.**

**Proposition 18.** i) $\mathcal{N}(y) = \emptyset$ *for* $y < 0$.

ii) $\mathcal{N}(y) = \mathcal{N}(2) = \{\nu \in BM(\mathbb{T}) : \langle \nu, 1\rangle = 1$ *and* $\nu \geq 0\}$ *for* $y \geq 2$.

iii) $\emptyset \neq \{\nu \in BM(\mathbb{T}) : d\nu(x) = 1 + \sum_{k=1}^{\infty} y_k \cos kx,\ \sum_{k=1}^{\infty} |y_k| \leq 1,\ y_k \leq 0\ (k \in \mathbb{N}_2)\} \subset \mathcal{N}(0) = \{\nu \in BM(\mathbb{T}) : \nu \geq 0,\ d\nu(x) = 1 + \sum_{k=1}^{\infty} y_k \cos kx,\ 1 + \sum_{k=1}^{\infty} y_k \geq 0,\ y_k \leq 0\ (k \in \mathbb{N}_2)\}$.

iv) *For any two values* $0 \leq y' < y'' \leq 2$ *we have* $\mathcal{N}(y') \subsetneq \mathcal{N}(y'')$.

v) *For all* $y \geq 0$ $\mathcal{N}(y)$ *is a convex, closed and bounded set in* $BM(\mathbb{T})$.

*Proof.* i) Suppose that $y < 0$ and $\nu \in \mathcal{N}(y)$. Consider the convolution $\nu * F_N = f_N \in \mathcal{T}_N$ for any $N \in \mathbb{N}$ where $F_N$ denotes the usual Fejér kernel. On one hand $f_N \geq 0$, on the other hand $f_N(0) = 1 + \left(1 - \frac{1}{N+1}\right) y_1 + \sum_{k=2}^{N} \left(1 - \frac{k}{N+1}\right) y_k \leq 1 + 2 + y \sum_{k=2}^{N} \left(1 - \frac{k}{N+1}\right)$. As the right-hand side tends to $-\infty$ with $N \to \infty$ by $y < 0$, we have proved i) by contradiction.

ii) For all $\nu \in BM(\mathbb{T})$, $\langle \nu, 1\rangle = 1$, $\nu \geq 0$ we have $\langle \nu, 2\cos kx\rangle \leq \langle \nu, 2\rangle = 2$.

iii) It suffices to prove the last equation, the others being easy consequences. Plainly the conditions on the right-hand side are exceeding the defining conditions

(105) for $\mathcal{N}(0)$ in two respect: by prescribing convergence of the Fourier representation, and by supposing

$$1 + \sum_{k=1}^{\infty} y_k \geq 0. \tag{127}$$

This last condition, together with $y_k \leq 0$ $(k \in \mathbb{N}_2)$, entails absolute convergence of the series (127), hence the Fourier representation must be absolutely uniformly convergent, and the measure $\nu$ is an absolutely continuous measure with a derivative having absolutely uniformly convergent series representation. The only thing to show that (127) holds for all $\nu \in \mathcal{N}(0)$. We can use the Fejér kernel $F_N$ and the convolution $F_N * \nu$, already used in part i) to get for arbitrary $N \in \mathbb{N}$

$$\begin{aligned} 0 \leq f_N(0) = (\nu * F_N)(0) &= 1 + \sum_{k=1}^{N} \left(1 - \frac{k}{N+1}\right) y_k \\ &\leq 1 + y_1 + \sum_{k=2}^{\infty} \left(1 - \frac{k}{N+1}\right) y_k. \end{aligned}$$

Using also $y_k \leq 0$ $(k \in \mathbb{N}_2)$, we can take limits with respect to $N \to \infty$, what yields (127).

iv) The inclusion is trivial. $\nu = \left(1 - \frac{y''}{2}\right) \cdot \lambda + \frac{y''}{2} \cdot \delta \in \mathcal{N}(y'')$ but $\nu \notin \mathcal{N}(y')$ shows that $\mathcal{N}(y') \neq \mathcal{N}(y'')$.

v)

$$\mathcal{N}(y) = \mathcal{N}(2) \cap \bigcap_{k=2}^{\infty} \left\{\nu \in BM(\mathbb{T}) : \langle 2\cos kx, \nu\rangle \leq y\right\}$$

$\mathcal{N}(2)$ is the intersection of a hyperplane defined by $\langle 1, \nu\rangle = 1$, and the (closed and convex) set of nonnegative measures. The other intersection is defined as the intersection of closed halfspaces. Therefore $\mathcal{N}(y)$ is convex and closed. Note that for any $\nu \in \mathcal{N}(y)$ $||\nu||_{BM(\mathbb{T})} = 2\pi$, cf. (109), hence $\mathcal{N}(y)$ is also bounded. □

**Proposition 19.** i) *For all $n \in \mathbb{N}$ there exists a unique $C_n$, $-2 \leq C_n < 0$, so that $\mathcal{N}_n(y) = \emptyset$ for $y < C_n$, but not for $y \geq C_n$.*

ii) *$C_n \nearrow 0$ as $n \to +\infty$.*

iii) *$\mathcal{N}_n(y) = \mathcal{N}_n(2) = \mathcal{N}(2)$ for $y \geq 2$.*

iv) *For any two values $C_n \leq y' < y'' \leq 2$ we have*

$$\mathcal{N}_n(y') \subsetneq \mathcal{N}_n(y'').$$

v) *For all $y \geq C_n$ $\mathcal{N}_n(y)$ is a convex, closed and bounded set in $BM(\mathbb{T})$.*

*Proof.* i) Plainly, as the inclusion part is trivial from statement iv), $C_n := \inf\{y : \exists \nu \in \mathcal{N}_n(y)\} = \sup\{y : \mathcal{N}_n(y) = \emptyset\}$.

First we prove $C_n < 0$, or more precisely, $C_n \leq -\frac{1}{n-1}$. To this end let us consider the function $g_n(x) := 1 - \frac{1}{n-1}\sum_{k=2}^{n}\cos kx$ and the measure $d\nu(x) = g_n(x)dx$. Plainly $\nu \in \mathcal{N}\left(-\frac{1}{n-1}\right)$ showing $C_n \leq -\frac{1}{n-1}$. On the other hand let $y < 0$ be arbitrary with $\mathcal{N}(y) \neq 0$ and let $\nu \in \mathcal{N}(y)$. If we choose $N = n$ in the construction of the Proof of Proposition 18 i), we obtain $0 \leq g_n(0) \leq 3 + y\left(\frac{N}{2} - 1 + \frac{1}{N+1}\right) \leq 3 - |y| \cdot \frac{n-2}{2}$. Thus $|y| \leq \frac{6}{n-2}$, proving also $C_n \to 0$ $(n \to +\infty)$. For small $n$ the above estimate can be substituted by the easier one

$$0 \leq \langle 1 + \cos 2x, \nu \rangle = 1 + \frac{y_2}{2} \leq 1 + \frac{y}{2},$$

showing $y \geq -2$; the measure $\nu_{\frac{\pi}{2}} := \frac{1}{2}\left(\delta_{\frac{\pi}{2}} + \delta_{-\frac{\pi}{2}}\right) \in \mathcal{N}_2(-2)$ shows that this estimate is sharp for $n = 2$.

We now prove $\mathcal{N}_n(C_n) \neq \emptyset$. This statement is a Cantor-type one, as $M_m := \mathcal{N}(C_n + \frac{1}{m})$ are closed (also convex) and nonempty sets of $BM(\mathbb{T})$ and plainly $\mathcal{N}_n(C_n) = \bigcap_{m=1}^{\infty} M_m$, where $M_m \supset M_M$ for all $m < M$.

To apply Cantor's Lemma, we only have to show that the sets $M_m$ are compact sets. That is not true in the original topology of $BM(\mathbb{T})$, but it holds true in the weak $*$ topology of $BM(\mathbb{T})$. Indeed, $\mathcal{N}_n(y)$ is bounded in view of (105) and (108), and all bounded sets of $BM(\mathbb{T})$ are conditionally compact in the weak $*$ topology. Moreover, $\mathcal{N}(y)$ is also closed, since by (114) $BM(\mathbb{T})_P$ is closed, and we have, using notation (110),

$$\mathcal{N}_n(y) = BM(\mathbb{T})_P \cap \left(\bigcap_{k=2}^{n} X(\cos kx, y)\right) \cap X(-1,-1) \cap X(1,1). \tag{128}$$

Thus Lemma 13 can be applied to show $\mathcal{N}_n(C_n) \neq \emptyset$.

ii) Monotonicity is obvious from definition, and $C_n \to 0$ is already proved.

iii) Follows from Proposition 18 ii) trivially.

iv) The inclusion is obvious. If $y'' > 0$, the example in Proposition 18 iv) fits here, too, showing $\mathcal{N}_n(y'') \setminus \mathcal{N}_n(y') \neq \emptyset$, while $\lambda \in \mathcal{N}_n(0)$ belongs to no $\mathcal{N}_n(y')$ with $y' < 0$. In case $y' < y'' < 0$ by the same way any $\nu' \in \mathcal{N}_n(y')$ with $\min_{2 \leq k \leq n} y_k(\nu') = y \leq y'$ can be a starting point to define $\nu'' = \frac{y''}{y} \cdot \nu' + \left(1 - \frac{y''}{y}\right)\lambda \in \mathcal{N}_n(y'')$ with $\nu'' \notin \mathcal{N}_n(y')$ since $\min_{2 \leq k \leq n} y_k(\nu'') = y'' > y'$.

v) Clear. $\square$

**Proposition 20.** *In the definitions* (107) *for $\beta$, $\vartheta$, $\beta_n$, $\vartheta_n$ $(n \in \mathbb{N})$, the supremum can be substituted by maximum since in case $\mathcal{N}(y) \neq \emptyset$, resp. $\mathcal{N}_n(y) \neq \emptyset$, the supremum is actually attained by some measure of $\mathcal{N}(y)$, resp. $\mathcal{N}_n(y)$.*

*Proof.* One proof can be given following the proof of $\mathcal{N}_n(C_n) \neq \emptyset$. However, it is easier to refer to the fact that $\mathcal{N}(y)$ and all $\mathcal{N}_n(y)$ $(n \in \mathbb{N})$ are closed sets, and thus the continuous functional $\nu \to \langle 2\cos x, \nu\rangle$ maps these sets to some closed sets of $\mathbb{R}$. (We may also note that the sets are convex and bounded, too, hence the image sets must be finite closed intervals.) Taking the supremums as in (107), we actually extremalize on these image sets of $\mathbb{R}$, and that concludes the argument. □

**Proposition 21.** i) $\vartheta_2(y) = \sqrt{2+y} \quad (C_2 = -2 \leq y \leq 2)$.

ii) $\beta_2(y) = \sqrt{2+y} \quad (C_2 = -2 \leq y \leq 2)$.

*Proof.* We already know that $C_2 = -2$ (cf. the end of the proof of part i) of Proposition 19.) Since $\beta_2(y) \geq 0$ and $\vartheta_2(y) \geq 0$ and both functions are nondecreasing in their domain of definition, it is enough to prove the statements for all $-2 < y \leq 2$. Let us fix one particular $y$, and let us choose two extremal measures $\mu, \nu \in \mathcal{N}_2(y)$ with Fourier series

$$d\mu(x) \sim 1 + \sum_{k=1}^{\infty} z_k \cos kx, \quad d\nu(x) \sim 1 + \sum_{k=1}^{\infty} y_k \cos kx \tag{129}$$

so that

$$z_1 = -\beta_2(y), \quad y_1 = \vartheta_2(y). \tag{130}$$

The extremal measures exist according to Proposition 20. Let us estimate $y_2$ and $z_2$ by the values of $y_1$ and $z_1$! That kind of estimation was already worked out in [22], Theorem 2.2 (see also the Remark after it). We get from this Theorem that

$$z_2 \geq 2\cos\left(2\arccos\left(\frac{-z_1}{2}\right)\right), \quad y_2 \geq 2\cos\left(2\arccos\left(\frac{y_1}{2}\right)\right). \tag{131}$$

Combining (129) and (130) with the inequalities $z_2 \leq y$, $y_2 \leq y$, coming from $\mu, \nu \in \mathcal{N}(y)$, we are led to

$$y \geq 2\cos\left(2\arccos\left(\frac{\beta_2(y)}{y}\right)\right), \quad y \geq 2\cos\left(2\arccos\left(\frac{\vartheta_2(y)}{2}\right)\right). \tag{132}$$

After some calculation this yields the estimates

$$\beta_2(y) \leq \sqrt{2+y}, \quad \vartheta_2(y) \leq \sqrt{2+y}. \tag{133}$$

Now we only have to show that this upper estimate is sharp. Let us consider the measure

$$\eta := \frac{1}{2}(\delta_w + \delta_{-w}) \quad \left(w := \arccos\left(\frac{z}{2}\right)\right), \quad \left(z := \sqrt{2+y}\right).$$

We have $\eta \geq 0$, and

$$d\eta(x) \sim 1 + \sum_{k=1}^{\infty} 2\cos(kw)\cos kx,$$

and thus the coefficient of $\cos x$ is $z$, and the coefficient of $\cos 2x$ is $2\cos(2w) = 2(2\cos^2(w) - 1) = z^2 - 2 = y$, verifying $\eta \in \mathcal{N}_2(y)$ and $\vartheta_2(y) \geq z = \sqrt{2+y}$. The translated measure $d\eta(x+\pi)$ shows by the same way $\beta_2(y) \geq \sqrt{2+y}$. These and (133) together concludes the proof of the Proposition. □

**Proposition 22.** i) *$\beta$ and $\vartheta$ are concave functions on $\mathcal{D}(\beta) = \mathcal{D}(\vartheta) = [0,\infty)$.*

ii) *For all $n \in \mathbb{N}$ $\beta_n$ and $\vartheta_n$ are concave functions on $\mathcal{D}(\beta_n) = \mathcal{D}(\vartheta_n) = [C_n,\infty)$.*

iii) *$\beta(y) = \vartheta(y) = \beta_n(y) = \vartheta_n(y) = 2$ for all $y \geq 2$ and $n \in \mathbb{N}$.*

iv) *For all $m > n$, $m,n \in \mathbb{N}$ and $y \geq C_m$, we have*

$$\beta_m(y) \leq \beta_n(y), \quad \vartheta_m(y) \leq \vartheta_n(y).$$

v) *For all $y \geq 0$ we have $\beta_n(y) \to \beta(y)$, $\vartheta_n(y) \to \vartheta(y)$ $(n \to \infty)$, uniformly in $y$.*

vi) *$\beta(0) = 1$, $\vartheta(0) \geq \frac{2}{\sqrt{3}}$.*

vii) *$\beta$ and $\vartheta$ are strictly increasing in $[0,2]$; $\beta_n$ and $\vartheta_n$ are strictly increasing in $[C_n,2]$ $(n \in \mathbb{N})$.*

*Proof.* i) Let $0 \leq y' \leq y \leq y''$ be arbitrary $y = \lambda y' + (1-\lambda)y''$ be the representation of $y$. Note that here we have $0 \leq \lambda \leq 1$. Suppose that $\nu' \in \mathcal{N}(y')$ and $\nu'' \in \mathcal{N}(y'')$ and consider the measure

$$\nu := \lambda\nu' + (1-\lambda)\nu'' \in BM(\mathbb{T}) \tag{134}$$

which is nonnegative as $\lambda \geq 0$ and $1-\lambda \geq 0$.

Plainly

$$\begin{aligned} \langle 1,\nu\rangle &= \lambda\langle 1,\nu'\rangle + (1-\lambda)\langle 1,\nu''\rangle = 1, \\ y_1 =& \langle 2\cos x,\nu\rangle = \lambda y_1' + (1-\lambda)y_1'', \\ y_k =& \langle 2\cos kx,\nu\rangle = \lambda y_k' + (1-\lambda)y_k'', \end{aligned} \tag{135}$$

where

$$\begin{aligned} d\nu(x) &\sim 1 + \sum_{k=1}^{\infty} y_k \cos kx, \\ d\nu'(x) &\sim 1 + \sum_{k=1}^{\infty} y_k' \cos kx, \\ d\nu''(x) &\sim 1 + \sum_{k=2}^{\infty} y_k'' \cos kx. \end{aligned} \tag{136}$$

Now the first and the third lines of (135) prove $\nu \in \mathcal{N}(y)$ as $y_k = \lambda y_k' + (1-\lambda)y_k'' \leq \lambda y' + (1-\lambda)y'' = y$ $(k \in \mathbb{N}_2)$. The second equation of (135) entails that

$$\begin{aligned} \vartheta(y) &= \sup\{y_1 : \nu \in \mathcal{N}(y)\} \\ &\geq \lambda \sup\{y_1' : \nu' \in \mathcal{N}(y')\} + (1-\lambda)\sup\{y_1'' : \nu'' \in \mathcal{N}(y'')\} \\ &= \lambda\vartheta(y') + (1-\lambda)\vartheta(y'') \end{aligned} \tag{137}$$

and similarly to (137) we also have

$$\beta(y) = \sup\{-y_1 : \nu \in \mathcal{N}(y)\} \geq \lambda\beta(y') + (1-\lambda)\beta(y''), \tag{138}$$

proving concavity.

ii) Similar to i).

iii) Proposition 18 ii) and Proposition 19 iii) entail that all functions are constant for $y \geq 2$. The same coefficient estimate

$$|\langle 2\cos kx, \nu\rangle| \leq \langle 2, \nu\rangle = 2 \quad (k \in \mathbb{N},\ \nu \in \mathcal{N}(2)), \tag{139}$$

already used in the proof of Proposition 19 iii), shows that $|y_1| \leq 2$. Now $\beta_n(2) = \beta(2) = 2$ is shown by $\delta_\pi \in \mathcal{N}(2)$, and $\vartheta_n(2) = \vartheta(2) = 2$ is shown by $\delta \in \mathcal{N}(2)$.

iv) Trivial in view of $C_n \leq C_m \leq y$ and $\emptyset \neq \mathcal{N}_m(y) \subseteq \mathcal{N}_n(y)$.

v) For any fixed particular $y \geq 0$ we have $\mathcal{N}(y) = \bigcap_{n=2}^\infty \mathcal{N}_n(y)$. Therefore $\beta(y) \leq \beta_n(y)$ $(n \in \mathbb{N})$ is trivial. To prove convergence of $\beta_n(y)$ to $\beta(y)$ at the point $y$, let us denote for all $n \in \mathbb{N}$

$$N_n := \Big\{\nu_n \in \mathcal{N}_n(y) : \langle 2\cos x, \nu_n\rangle \leq -\beta_n(y)\Big\} = \mathcal{N}_n(y) \cap X\big(2\cos x, \beta_n(y)\big). \tag{140}$$

Note that in view of Proposition 20 $N_n \neq \emptyset$, and the sets $N_n$ satisfy all the conditions of Lemma 13 in view of (128) and (140). Hence $N = \bigcap\limits_{n=2}^{\infty} N_n$ is nonempty. One can easily see that any $\nu \in N$ belongs to $\mathcal{N}(y)$ and

$$\langle 2\cos x, \nu\rangle \leq -\lim_{n\to\infty} \beta_n(y), \tag{141}$$

proving $\beta(y) \geq \lim\limits_{n\to\infty} \beta_n(y)$. Now we have $\beta_n(y) \to \beta(y)$ monotonically nonincreasingly in the pointwise sense on the whole $[0,2]$. But for the concave and hence continuous functions $\beta_n$ and $\beta$ that entails also uniform convergence on $[0,2]$ by Dini's monotone convergence criteria (cf. e.g. [8], (7.2.2), p. 129). With part iii) that settles uniform convergence, too. A similar argument works for $\vartheta$ as well.

vi) The easy examples $d\nu_+(x) = (1+\cos x)dx$, $d\nu_-(x) = (1-\cos x)dx$ show that $\vartheta(0) \geq 1$ and $\beta(0) \geq 1$. To show that $\vartheta(0) > 1$, one may consider the trigonometric polynomial

$$h(x) = 1 + \frac{1}{\cos\frac{\pi}{6}}\cos x - \frac{\tan\frac{\pi}{6}}{3}\cos 3x = 1 + \frac{2}{\sqrt{3}}\cos x - \frac{1}{3\sqrt{3}}\cos 3x \tag{142}$$

and the corresponding measure $d\nu(x) = h(x)dx$. The only thing to check is $h \geq 0$, which can be done directly, or we can refer to the $k = 3$ case of Proposition 2.1 of [22]. On the other hand, Proposition 18 iii) entails that for $\nu \in \mathcal{N}(0)$ we can not have $\langle 2\cos x, \nu\rangle = y_1 < -1$, and this proves $\beta(0) = 1$.

vii) Follows from parts i), ii), iii) and Proposition 21. □

**11.5.**

**Proposition 23.** i) *For $1 \leq a < 2$ we have*

$$\omega(a) = a \cdot \max_{y>0} \frac{\beta(y) - 2/a}{y} + a + 1. \tag{143}$$

ii) *For $a \leq -1$, $a \in \mathcal{D}(\omega)$ we have*

$$\omega(a) = (-a) \max_{y>0} \frac{\vartheta(y) + 2/a}{y} + a + 1. \tag{144}$$

**Remark 3.** For $-1 \leq a \leq 1$ we have $\omega(a) = \alpha(a) = 1 + a$ according to Propositions 14 i) and Theorem 4 i). Also for $-1 \leq a \leq 1$

$$\sup_{y>0} \frac{a \cdot \beta(y) - 2}{y} = \lim_{y\to+\infty} \frac{a \cdot \beta(y) - 2}{y} = 0 \tag{145}$$

and

$$\sup_{y>0} \frac{(-a)\vartheta(y) - 2}{y} = \lim_{y\to+\infty} \frac{(-a)\vartheta(y) - 2}{y} = 0, \tag{146}$$

since $|a| \leq 1$, and $0 \leq \beta(y) \leq 2$, $0 \leq \vartheta(y) \leq 2$ (see Proposition 22) and hence the numerators of these functions are always nonpositive. In this sense the statement is valid for all $a$, but to emphasize the given forms, where sup is changed to max and $a$ has been brought out, we used the above formulation.

*Proof.* i) For $a = 1$ the statement is trivial according to the above Remark. For $a > 1$ let us take two extremal measures $\tau \in \mathcal{M}(a)$ and $\nu_0 \in \mathcal{N}(y_0)$ with $\max_{y>0} \frac{\beta(y)-2/a}{y} = \frac{\beta(y_0)-2/a}{y_0}$. Since $a > 1$ and $\beta(0) = 1$ we see that $\varphi_a(y) = \varphi(y) = \frac{\beta(y)-2/a}{y}$ is negative for small $y > 0$, while for $y = 2$ $\varphi(2) = 1 - 1/a > 0$, and for $y > 2$ $\varphi(y) < \varphi(2)$ and $\varphi(y) \to 0$ $(y \to +\infty)$. Hence there exists a $y_0 \notin (0, 2]$, depending on $a$, where $\varphi_a(y_0)$ is a maximum. Also $\tau$ and $\nu_0$ must exist in view of Proposition 20 and Theorem 4 i). First, we recall

$$\tau + \delta \geq t \cdot \lambda; \quad t = \omega(a) \tag{147}$$

and define

$$\mu := \tau + \delta - t \cdot \lambda \geq 0, \tag{148}$$

where with the Fourier expansion in (103) and (104) we are led to

$$d\mu(x) \sim (b+1-t) + \left(2 - \frac{2b}{a}\right)\cos x + \sum_{k=2}^{\infty}(2+t_k)\cos kx, \qquad (149)$$
$$b \in \mathbb{R},\ t_k \le 0\ (k \in \mathbb{N}_2).$$

Now by $\mu \geq 0$ we have also $b+1-t \geq 0$. In case of $b+1-t=0$ the trivial argument (109) would give $\mu \equiv 0$, leading in view of the coefficient of $\cos x$ to the equation $b=a$ and thus $a+1-t=0$. But $t=\omega(a)=\alpha(a)\geq 2a > 1+a$ for $a>1$ according to Propositions 14 ii) and 3.5 i), thus excluding $b+1-t=0$ for $a>1$. We get

$$b+1-t>0 \quad \text{if} \quad a>1, \qquad (150)$$

and we can introduce the new normalized measure

$$\nu := \frac{1}{b+1-t}\mu, \quad d\nu(x) \sim 1 + \frac{2(a-b)}{a(b+1-t)}\cos x + \sum_{k=2}^{\infty}\frac{2+t_k}{b+1-t}\cos kx. \qquad (151)$$

Denoting

$$y_1 := -\frac{2(b-a)}{a(b+1-t)}, \quad y_k := -\frac{2+t_k}{b+1-t} \quad (k\in\mathbb{N}_2), \quad y := \frac{2}{a(b+1-t)}, \qquad (152)$$

we immediately get that $\nu \in \mathcal{N}(y)$ with the parameters and coefficients in (152). Consequently, we have by definition

$$\frac{2(b-a)}{a(b+1-t)} \leq \beta(y) \quad \left(y - \frac{2}{b+1-t} > 0\right). \qquad (153)$$

Let us use the definition of $y$ and $t=\omega(a)$ in the left-hand side to express (153) by $y$ and $\omega(a)$ as

$$\frac{y}{a}\left\{\left(\frac{2}{y} - 1 + \omega(a)\right) - a\right\} \leq \beta(y), \qquad (154)$$

or, after some calculation,

$$\omega(a) \leq a \cdot \frac{\beta(y) - 2/a}{y} + a + 1. \qquad (155)$$

That proves that the left-hand side of (143) can not exceed the right-hand side. Next we start by considering the extremal measure $\nu_0 \in \mathcal{N}(y_0)$ and define

$$b := a\left(\frac{\beta(y_0)}{y_0}+1\right), \quad t := \frac{-2}{y_0}+1+b, \quad t_k := \frac{2(y_k-y_0)}{y_0} \quad (k\in\mathbb{N}_2),$$
$$\tau_0 := \frac{2}{y_0}\nu - \delta + t\cdot\lambda \in BM(\mathbb{T}), \quad d\tau_0(x) \sim b + \left(\frac{2}{y_0}y_1 - 2\right)\cos x + \sum_{k=2}^{\infty} t_k \cos kx. \qquad (156)$$

We immediately have $t_k \le 0$ $(k \in \mathbb{N}_2)$ and from the extremality of $\nu_0 \in \mathcal{N}(y_0)$ we also have $y_1 = -\beta(y_0)$. Moreover, in view of the definition of $b$, we have for $t_1$, the coefficient of $\cos x$ in the Fourier expansion of $\tau_0$, the equation

$$t_1 = \frac{-2\beta(y_0)}{y_0} - 2 = b\left(-\frac{2}{a}\right). \tag{157}$$

Now (156)–(157) yield $\tau_0 \in \mathcal{M}(a)$, and, as $\nu_0 \in \mathcal{N}(y_0)$ entails $\nu_0 \ge 0$, we immediately get $\tau_0 + \delta \ge t \cdot \lambda$ proving that

$$\omega(a) \ge t. \tag{158}$$

Now let us substitute the parameters (156) in (158) to obtain

$$\begin{aligned}\omega(a) &\ge -\frac{2}{y_0} + 1 + a\left(\frac{\beta(y_0)}{y_0} + 1\right)\\ &= a \cdot \frac{\beta(y_0) - 2/a}{y_0} + a + 1.\end{aligned} \tag{159}$$

Comparing (155) and (159) proves the assertion.

ii) The proof is very similar to i), hence we omit a few details and give here only the main steps and formulas. Again we suppose $a < -1$, and check that $\psi_a(y) := \psi(y) := \frac{\vartheta(y)+2/a}{y}$ has a positive maximum attained for some $y$ in $0 < y \le 2$. The $\le$ part will be proved by taking an extremal $\tau \in \mathcal{M}(a)$ and following the preceding argument from (145) up to (152) with the only alteration that here in place of (150) we have

$$b + 1 - t > 0 \quad \text{if} \quad a < -1 \tag{160}$$

because of the relations $t = \omega(a) \ge 0 > 1 + a$ $(a < -1)$. Now in place of (153) we will obtain from the extremality of $\tau$ that

$$\frac{2(b-a)}{(-a)(b+1-t)} \le \vartheta(y) \qquad \left(y = \frac{2}{b+1-t} > 0\right), \tag{161}$$

and similarly to (154)–(156), some calculation leads to

$$\omega(a) = t \le (-a)\frac{\vartheta(y) + 2/a}{y} + a + 1. \tag{162}$$

The converse direction goes like (156)–(158) with the only change that here we take $y_1 = \vartheta(y_0)$ in place of $-\beta(y_0)$. Hence the same change occurs in (159) and we get the $\ge$ part. $\square$

**Remark 4.** We have to note here that implicitly we used that $\vartheta(y)+2/a$ is positive only for $y > 0$, i.e. $\vartheta(0) \le 2/|a|$. Now we really have

$$\vartheta(0) = \frac{2}{-A}, \tag{163}$$

a duality-type relation between different extremal problems, cf. [21], in particular the discussion around (116)–(121).

Note that this settles the existence of maximum for $\psi_a(y)$ in $y > 0$ for all $a > A$, but leaves the question open if $A \in \mathcal{D}(\alpha)$ and $a = A$. In this case for small $y$ $(\vartheta(y) + 2/A)/y$ has a small, but positive numerator and the denominator is also positive. Thus we can extend $\psi_A$ to 0 as

$$\begin{aligned}\psi_A(0) &= (-A) \lim_{y\to 0+} \frac{\vartheta(y) + 2/A}{y} + A + 1 = (-A) \lim_{y\to 0+} \frac{\vartheta(y) - \vartheta(0)}{y} + A + 1 \\ &= (-A)\vartheta'(0+) + A + 1\end{aligned} \tag{164}$$

in case it is finite. In turn, if (164) is finite, by concavity we conclude that the maximum is attained at 0, and we conclude

$$\alpha(A) = \omega(A) = (-A)\vartheta'(0+) + A + 1. \tag{165}$$

On the other hand, if $\vartheta'(0+) = +\infty$, similarly to (159) it is easy to show that we will have $\lim\limits_{a\to A+} \alpha(a) = +\infty$, and hence $A \notin \mathcal{D}(\alpha)$.

Similarly, from $\alpha(a) \to +\infty$ $(a \to 2-)$ we can conclude that

$$\beta'(0+) = +\infty. \tag{166}$$

Later even the asymptotic order of $\beta$ will be specified, so we leave this question for the moment.

Let us point out the geometric interpretation of the maximum in (143). The concave curve $\{(y, \beta(y)) : y \geq 0\}$ defines a convex domain of points lying below the curve. The maximum is just the slope of one of the tangent straight lines drawn from the outer point $(0, 2/a)$ to this convex domain. (The other tangent is just the second coordinate axis.)

**Proposition 24.** *We have for all $n \in \mathbb{N}$ the relations*

i) $\beta_n(0) = \frac{2}{B_n}$;

ii) $\vartheta_n(0) = \frac{-2}{A_n}$;

iii) *For all $a \in [0, B_n]$*

$$\omega_n(a) = a \sup_{y>0} \frac{\beta_n(y) - 2/a}{y} + a + 1;$$

*in particular,*

$$\omega_n(B_n) = B_n \cdot \beta_n'(0+) + B_n + 1;$$

iv) *For all* $a \in [A_n, 0]$

$$\omega_n(a) = (-a) \sup_{y>0} \frac{\vartheta_n(y) + 2/a}{y} + a + 1;$$

*in particular*

$$\omega_n(A_n) = (-A_n) \cdot \vartheta_n'(0+) + A_n + 1.$$

*Proof.* Follows similarly to the argument of Proposition 23 and the Remark after it. We omit the details. $\square$

Next we define another extremal quantity as follows.

$$Z := \inf \Big\{ y > 0 : \exists \zeta \in BM(\mathbb{T}), \quad \zeta \geq 0, \quad z \geq 2(1-y), \tag{167}$$
$$d\zeta(x) \sim 1 - z \cos x + \sum_{k=2}^{\infty} z_k \cos kx, \quad z_k \leq y \quad (k \in \mathbb{N}_2) \Big\}.$$

**Proposition 25.** i) *There exists a unique point* $y_U$ *in* $(0, 2]$ *so that*

$$\beta(y_U) = 2(1 - y_U).$$

ii) *For the point* $y_U$ *we have* $Z = y_U$.

iii) *We have* $\Omega = \frac{2}{y_U} - 1 = \frac{2}{Z} - 1$.

*Proof.* i) The functions $\beta(y)$ and $2(1-y)$ are continuous and strictly monotonous in the opposite direction from 0 to 2 and from 2 to $-2$ in the domain $[0, 2]$. Hence there exists a unique solution of the equation $\beta(y) = 2(1 - y)$ in the interval $(0, 2)$.

ii) Denote the set of measures used in the definition of $Z$ as $\mathcal{Z}(y)$. Then

$$Z := \inf \big\{ y > 0 : \mathcal{Z}(y) \neq \emptyset \big\}. \tag{168}$$

Now if $y > Z$, we have $\mathcal{Z}(y) \neq \emptyset$, and, as $\mathcal{Z}(y) \subset \mathcal{N}(y)$, we find that $\beta(y) = \max\{-\langle 2\cos x, \zeta\rangle : \zeta \in \mathcal{Z}(y)\} \geq 2(1-y)$. Hence, in view of the definition of $y_U$ and the monotonicity of $\beta(y)$ and $2(1-y)$, we conclude $y \geq y_U$ and a fortiori $Z \geq y_U$. Conversely, if $y > y_U$, then $\beta(y) > \beta(y_U) = \max\{\langle -2\cos x, \zeta\rangle : \zeta \in \mathcal{N}(y_U)\}$, and for any extremal measure $\zeta_0 \in \mathcal{N}(y_U)$, we have $\langle -2\cos x, \zeta_0\rangle = \beta(y_U) = 2(1 - y_U)$, hence $\zeta_0 \in \mathcal{Z}(y_U)$ and $Z \leq y_U$.

iii) Let $\mathcal{K}$ be the measure set in (99) where the defining supremum for $\Omega$ is defined. Note that $\mathcal{K}$ contains $\delta_\pi$, hence $\mathcal{K} \neq \emptyset$. Moreover we have for any $\kappa \in \mathcal{K}$

$$\begin{aligned} 0 \leq ||\kappa||_{BM(\mathbb{T})} = \int |d\kappa| = \int d\kappa = 2\pi \cdot b_0 \leq 2\pi(2 - b_1) = \\ = 2\pi\big(1 + (1 - b_1)\big) \leq 2\pi(1 + \Omega), \end{aligned} \tag{169}$$

and using the estimate $\Omega \le U$ (stated already in (100) as a result implicitly contained already in [31], and proven in Corollary 4 below) we immediately get that $\mathcal{K}$ is bounded. Note that $\mathcal{K}$ is also closed and convex, and can be represented in the form of the intersection of a set of closed halfspaces generated by functionals from $C(\mathbb{T})$, hence $\mathcal{K}$ is also weakly $*$ compact and the sup in (99) is actually a maximum. Now for the extremal measure $\kappa \in \mathcal{K}$ we consider its Fourier series (97) and prove that $b_0 > 1$ and $b_0 + b_1 = 2$ for $\kappa$. Indeed, in case $b_0 \le 1$ we must have $|b_1| \le 1$, $\Omega = 1 - b_1 \le 2$, and the known examples are much better than that. Also if $b_0 + b_1 < 2$, one can consider $d\kappa^*(x) = d\kappa(x) + \frac{2}{3}(2 - b_0 - b_1) \cdot (1 - \cos x)dx$, $b_0^* + b_1^* = \langle 1 - \cos x, \kappa^* \rangle = b_0 + b_1 + \frac{2}{3}(2 - b_0 - b_1)\left(1 + \frac{1}{2}\right) = 2$, hence $\kappa^* \in \mathcal{K}$, and $b_1^* < b_1$ would provide a contradiction. $\square$

Now let us define the measure

$$\nu = \frac{1}{b_0} \cdot \kappa \ge 0.$$

Plainly $\nu \in \mathcal{N}\left(\frac{2}{b_0}\right)$, hence $\beta\left(\frac{2}{b_0}\right) \ge \frac{-2b_1}{b_0} = \frac{2(b_0-2)}{b_0} = 2\left(1 - \frac{2}{b_0}\right)$. Let $y_0$ be $\frac{2}{b_0}$, then we see $\beta(y_0) \ge 2(1 - y_0)$, hence $y_0 \ge y_U$. From this we get $b_0 = \frac{2}{y_0} \le \frac{2}{y_U}$, hence $\Omega = 1 - b_1 = b_0 - 1 \le \frac{2}{y_U} - 1$. Similarly, for $y_U$ we can take any $\beta$-extremal measure $\nu \in \mathcal{N}(y_U)$ and consider the measure

$$\kappa = \frac{2}{y_U}\nu \in \mathcal{K}$$

proving $\Omega \ge 1 - b_1 = 1 + \frac{2}{y_U} \cdot \frac{\beta(y_U)}{2} = 1 + \frac{2}{y_U}(1 - y_U) = \frac{2}{y_U} - 1$.

Putting $[2, n]$ in place of $\mathbb{N}_2$ one can also introduce $\mathcal{Z}_n$ and $\mathcal{K}_n$, and the corresponding extremal quantities $Z_n$ and $\Omega_n$ $(n \in \mathbb{N}_2)$. It is no surprise now that we have the analogous

**Proposition 26.** *For arbitrary $n \in \mathbb{N}_2$ the following statements hold true.*

i) *There exists a unique point $y_{U,n}$ in $(0, 2]$ so that $\beta_n(y_{U,n}) = 2(1 - y_{U,n})$.*

ii) *For the point $y_{U,n}$ we have $Z_n = y_{U,n}$.*

iii) *We have $\Omega_n = \frac{2}{y_{U,n}} - 1 = \frac{2}{Z_n} - 1$.*

iv) *$Z_n \to Z$ monotonically increasingly, and $\Omega_n \to \Omega$ a nonincreasing way.*

**Corollary 4.** *We have $U = \Omega$ and also $U_n = \Omega_n$ $(n \in \mathbb{N}_2)$.*

*Proof.* As the proofs are very similar, we prove only $U = \Omega$. The easy part is $U \ge \Omega$, essentially already proved by van der Waerden [31] the idea dating back to

Landau [18]. Indeed, let $f \in \mathcal{F}(a)$ and $\kappa \in \mathcal{K}$ be any particular elements, we then have by $f \geq 0$, $\kappa \geq 0$ and using $b_k \leq 1$, $a_k \geq 0$ ($k \in \mathbb{N}_2$) that

$$\begin{aligned} 0 \leq \langle f, \kappa \rangle = b_0 + ab_1 + \sum_{k=2}^{\infty} a_k b_k \leq (b_0 - 1) + a(b_1 - 1) \\ + \left(1 + a + \sum_{b=2}^{\infty} a_k\right) = b_0 - 1 + a(b_1 - 1) + f(0). \end{aligned} \tag{170}$$

Now let us also apply $b_0 + b_1 \leq 2$ for $\kappa \in \mathcal{K}$, and get for $a > 1$ the inequalities

$$(1 - b_1)(a - 1) \leq b_1 - 1 + b_0 - 1 + f(0) \leq f(0). \tag{171}$$

Now let us take supremum over $\mathcal{K}$ at the left, and infimum over $\mathcal{F}(a)$ at the right-hand side to get

$$\Omega(a - 1) \leq \alpha(a). \tag{172}$$

Dividing by $a - 1$ ($> 0$) and minimizing $U(a) = \frac{\alpha(a)}{a-1}$, we get $\Omega \leq U$. (Note that the minimum place is $a = a_U$, cf. Proposition 17 iv) for the definition and uniqueness.)

Now let us prove the converse! We start with noting that by Proposition 25 iii) $\Omega = \frac{2}{y_U} - 1$, and choose $a = a_\Omega$ such that the maximum at the right-hand side of (143) is attained at $y = y_U$. Note that $0 < y_U < 0.5$ is trivial, and for (any one of the) tangential lines of $\beta$ at the point $(y_U, \beta(y_U))$ the intersection point of the straight line with the second coordinate axis defines such an $a_\Omega$ by $\left(0, \frac{2}{a}\right)$ being the intersection point. Hence we conclude the existence of such an $a_\Omega$. Consequently, with $a = a_\Omega$ and using the Duality Theorem (Theorem 4 i)), we get

$$\begin{aligned} U \leq \frac{\alpha(a_\Omega)}{a_\Omega - 1} = \frac{\omega(a_\Omega)}{a_\Omega - 1} = \frac{1}{a_\Omega - 1}\left(a_\Omega \frac{\beta(y_U) - 2/a_\Omega}{y_U} + a_\Omega + 1\right) \\ = \frac{1}{a-1}\left(a\frac{2(1-y)}{y} - \frac{2}{y} + a + 1\right) = \frac{1}{a-1}\left((a-1)\frac{2}{y} - a + 1\right) = \frac{2}{y_U} - 1 = \Omega. \end{aligned} \tag{173}$$

□

We may note that the above mentioned duality relation enables us to give another form of $\omega(a)$, which has the interesting feature that only the goal function to be maximalized is dependent on $a$, but not the set of measures on what the maximization takes place. Namely, we have

$$\alpha(a) = \omega(a) = \sup\left\{(1-b_1)(a-1)+(2-b_0-b_1) : \exists\kappa \in \mathcal{K} \quad (\text{with } (97)\text{–}(98))\right\}. \tag{174}$$

Indeed let us define the right-hand side as $\zeta(a)$, and define also the auxiliary quantity

$$w(b) := \sup\left\{1 - b_1 : \exists\kappa \in \mathcal{K}, \quad b_0 = b \quad (\text{with } (97)\text{–}(98))\right\}. \tag{175}$$

Plainly with $y := \frac{2}{b}$ the function $w(b)$ is related to $\beta(y)$ as

$$w(b) = 1 + \frac{b}{2}\beta(y) \quad \left(y := \frac{2}{b}\right), \tag{176}$$

since for $\kappa \in \mathcal{K}$ with $b_0 = b$ the measure $\nu := \frac{1}{b}\kappa \in \mathcal{N}(y)$, and for $\nu \in \mathcal{N}(y)$ the measure $\kappa := \frac{2}{y}\nu \in \mathcal{K}$. Plainly

$$\begin{aligned}
\zeta(a) &= \sup\big\{(1-b_1)a + (1-b_0) : \exists\kappa \in \mathcal{K} \quad (\text{with (97)–(99)})\big\} = \\
&= \sup\big\{(1-b_0) + aw(b_0) : \kappa \in \mathcal{K}\big\} = \sup_{b_0>0}\left\{1 - b_0 + a\left(1 + \frac{b_0}{2}\beta\left(\frac{2}{b_0}\right)\right)\right\} \\
&= \sup_{y>0}\left\{a\left(1 + \frac{\beta(y)}{y}\right) + 1 - \frac{2}{y}\right\} = a\sup_{y>0}\frac{\beta(y) - 2/a}{y} + a + 1 = \omega(a)
\end{aligned}$$

by (144), Proposition 23 i).

## §12. Concluding remarks and further questions for research

One can ask if Landau's extremal problems are interesting even now. We have already mentioned that they can be of interest for practical applications, in particular for computational number theory, as in [26]. Let us mention a further point and comment connections to recent publications like [9].

Already Landau proved the Prime Ideal Theorem, and later on further generalizations of Dirichlet's and Riemann's approach (use of multiplicative generating functions, i.e. Dirichlet series, in the study of multiplicative problems) appeared. A quite general setup is the *Beurling theory of prime distribution.* Now for Beurling primes the de la Vallée Poussin – Landau method works, but no other refined techniques can be utilized, since there are counterexamples: Diamond, Montgomery and Vorhauer [9] has constructed recently a Beurling set of primes so that no better zero-free region of $\zeta(s)$, and no better error term of $\pi(x)$, can be established than $\sigma > 1 - C/\log t$ and $xe^{-c\sqrt{\log x}}$, respectively.

Also, these problems are related to, or similar to, and have common generalizations with many other important families of extremal problems. So we are convinced that further research of them has merit not only for the analytical beauty and difficulty of them. Hence let us end this work by listing a few questions.

1.) The asymptotic order of $\alpha(a)$ when $a \to 2 - 0$ was determined in [22]. It is of interest to obtain more precise descriptions of values of $\alpha(a)$, in particular when $a \to 2 - 0$.

2.) We have seen that the extremal function in the $\alpha$-problem is a polynomial when say $-3/4 < a < \sqrt{2}$. (We can calculate this a bit further.) Do we have for all $a \in \mathcal{D}(\alpha)$ that there is an $N := N(a)$ so that $\alpha(a) = \alpha_N(a)$? (If so, the "right" (minimal) degree $N(a) \to \infty$ when $a \to 2 - 0$.) Having $N(1.85)$, say, would allow to *exactly* determine $U$, $V$, $W$ by finite range computer search.

3.) We have seen (Chakalov!) that sometimes $U_{k+1} = U_k$. It seems that in the dual (van der Waerden-type) extremal problem $\omega(a)$ and $\Omega$ for measures, we have vanishing Fourier coefficients for exactly those indices $k \in \mathbb{N}$. Prove or disprove!

4.) Determine $A$ and $A_n$ (left endpoints of $\mathcal{D}(\alpha)$ and of $\mathcal{D}(\alpha_n)$, resp.). These lead to extremal problems in themselves: for how large an $a$ can an even Fourier series $g(x)$ (a cosine polynomial of degree $n$) with $\widehat{g}(0) = 1$ and $\widehat{g}(1) = 0$ be strictly positive definite while $g(x) + a\cos x \geq 0$?

5.) It is possible to consider similar, however not only positive definite, but *signed* Landau-problems, i.e. instead of $a_k \geq 0$ we can assume arbitrary sign conditions on various $k$'s. We already have the duality [21]. Note that considering negative $a$ leads to this question naturally.

6.) Let $G$ be a locally compact Abelian group. Develop the similar theory.

We have emphasized several times that the Landau extremal problems are related to many classical and current extremal problems. Let us give an example for the Landau problem with some sign conditions.

**Example 1.** Assume, as sign condition, that $\widehat{f} = 0$ outside $D$, where $D \subset \widehat{G}$ is a domain in the dual group. (E.g. if $G = \mathbb{T}$ and $\widehat{G} = \mathbb{Z}$, then considering polynomials of degree $\leq n$ is equivalent to assume $\widehat{f} = 0$ outside $[-n, n]$.) Then the problem is about the minimal value of $f(0)$ when $a = a_1 = 2\widehat{f}(1)$ is given.

Changing the role of $G$ and $\widehat{G}$ and taking $\varphi := \widehat{f}$, we obtain the following extremal problem:

i) $\varphi = 0$ on $G \setminus D$, i.e. $\varphi$ is supported in $D$;

ii) $\widehat{\varphi} \geq 0$, i.e. $\varphi$ is positive definite;

iii) $\varphi(0) = 1$ (normalization);

iv) $\varphi(1) = a/2$;

and then we seek to minimize $\int \varphi = \widehat{\varphi}(0)$.

If one only looks for the largest possible value of $a$ so that the problem has a finite solution, (e.g. if we look for $B_n$), then the extremal problem becomes maximization of $\varphi(1)$ under the conditions (i)–(iii) given. This is called "pointwise Turán problem" (although in $\mathbb{R}$ was already considered by Boas and Kac [3] in the forties). See [15] and the references therein. In the special case of $\alpha_n$, it is an extremal problem solved by Fejér [11] and Szász [28] – that is the exact value of $B_n$ given above.

## References


[1] Arestov, V. V., On an extremal problem for nonnegative trigonometric polynomials, *Trudy Inst. Math. Mech. Ukr. Acad. Nauk* **1** (1992), 50–70 (in Russian).

[2] Arestov, V. V. and Kondrat'ev, V. P., On an extremal problem for nonnegative trigonometric polynomials, *Mat. Zametki* **47** (1990), No 1, 15–28 (in Russian, translated as Mathematical Notes).

[3] Boas, R. P. Jr. and Kac, M., Inequalities for Fourier transforms of positive functions, *Duke Math. J.* **12** (1945), 189–206.

[4] Chakalov, L. / Tschakaloff, L., Trigonometrische Polynome mit einer minimumeigenschaft, *Jahrbuch der Universität Sofia, Phys.-Math. Fakultät* **19** (1923), 355–387 (in Bulgarian).

[5] Chakalov, L. / Tschakaloff, L., Trigonometrische Polynome mit einer Minimumeigenschaft, *Annali di Scuola Normale Superiore di Pisa (2)* **9** (1940), 13–26.

[6] Chandrasekharan, K., *Arithmetical Functions*, Springer, Berlin – Heidelberg – New York, 1970.

[7] Compson, E. T., *Asymptotitc expansions*, Cambridge University Press, 1965.

[8] Dieudonné, J., *Foundations of modern analysis*, Academic Press, New York – London, 1960.

[9] Diamond, H. G., Montgomery, H. L. and Vorhauer, U. M., Beurling primes with large oscillation, *Math. Ann.* **334** (2006) 1, 1–36.

[10] Edwards, R. E., *Functional Analysis*, Holt, Rinehart and Winston, New York – Toronto – London, 1965.

[11] Fejér, L., Über trigonometrische Polynome, *J. angew. Math.* **146** (1915), 53–82. See also in Fejér, L., *Gesammelte Arbeiten*, Bd. I., 842–872, Akadémiai Kiadó, Budapest, 1970.

[12] Ford, K., Zero-free regions for the Riemann zeta function, *Number theory for the millennium, II* (Urbana, IL, 2000), 25–56.

[13] French, S. H., Trigonometric polynomials in prime number theory, *Illinois Journal of Mathematics* **10** (1966), 2, 240–248.

[14] Heathbrown, D. R., Zero-free regions for Dirichlet $L$-functions and the least prime in an arithmetic progression, *Proc. London Math. Soc.* **64** (1992), 265–338.

[15] Kolountzakis, M. N. and Révész, Sz. Gy., On pointwise estimates of positive definite functions with given support, *Canad. J. Math.* **58** (2006), 2, 401–418.

[16] Kondrat'ev, V. P., On some extremal properties of positive trigonometric polynomials, *Mat. Zametki* **22** (1977), No 3, 371–374 (in Russian, translated as Mathematical Notes).

[17] Landau, E., Beiträge zur analitische Zahlentheorie, *Rendiconti del Circolo Matematico di Palermo* **26** (1908), 169–302.

[18] Landau, E., *Handbuch der Lehre von der Verteilung der Primzahlen*, Teubner, Leipzig – Berlin, 1909.

[19] Landau, E., Eine Frage über trigonometrische Polynome, *Annali di Scuola Normale Superiore di Pisa (2)* **2** (1933), 209–210.

[20] Landau, E., Nachtrag zu meiner Arbeit "Eine Frage über trigonometrische Polynome", *Annali di Scuola Normale Superiore di Pisa (2)* **5** (1936), 141.

[21] Révész, Sz. Gy., The least possible value at zero of some nonnegative cosine polynomials and equivalent dual problems, *Journal of Fourier Analysis and Applications, Kahane Special Issue* (1995), 485–508.

[22] Révész, Sz. Gy., Minimization of maxima of nonnegative and positive definite cosine polynomials with prescribed first coefficients, *Acta Sci. Math. (Szeged)* **60** (1995), 589–608.

[23] Révész, Sz., On some extremal problems of Landau, *Serdica Math. J.* **33** (2007), 125–162.

[24] Reztsov, A. V., Certain extremal properties of nonnegative trigonometric polynomials, *Mat. Zametki* **39** (1986), No 2, 245–252 (in Russian, translated as Mathematical Notes).

[25] Rockafellar, R. T., *Convex Analysis*, Reprint of the 1970 original, Princeton University Press, Princeton, 1997.

[26] Rosser, J. B. and Schoenfeld, L., Approximate formulas for some functions of prime numbers, *Illinois J. Math.* **6** (1962), 64–92.

[27] Rosser, J. B. and Schoenfeld, L., Sharper Bounds for the Chebyshev Functions $\vartheta(x)$ and $\psi(x)$, *Mathematics of Computation* **29** (1975), No 129, 243–269.

[28] Szász, O., Elementare Extremalprobleme über nichtnegative trigonometrische Polynome, *Sitzungsberichte Bayerische Akad. der Wiss. Math.-Phys. Kl.* (1927), 185–196. See also in Szász, O. *Collected Mathematical Papers*, University of Cincinnati, 1955, pp. 734–735.

[29] Stechkin, S. B., On some extremal properties of nonnegative trigonometric polynomials, *Mat. Zametki* **7** (1970), No 4, 411–422 (in Russian).

[30] (de) la Vallée Poussin, Ch.-J. de, Sur la function $\zeta(s)$ de ..., *Memoire couronnés ... de Belgique* **59** (1899), No 1, 1–74.

[31] van der Waerden, B. L., Über Landau's Beweis des Primzahlsatzes, *Math.–Zeitschrift* **52** (1949), 649–653.

[32] Walfisz, A., *Weylsche Exponentialsummen in der neueren Zahlentheorie*, Berlin, 1963.

[33] Westphal, H. Über die Nullstellen der Riemannschen Zetafunction im kritischen Streifen, *Schriften des Math. Seminars und des Instituts für angewandte Math. der Universität Berlin* **4** (1938), 1–31.