INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International
300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND
WEISS, ALFRED REINHOLD

THE LEAST PRIME IDEAL WITH PRESCRIBED DECOMPOSITION BEHAVIOUR

The Ohio State University

University Microfilms International

300 N. Zeeb Road, Ann Arbor, MI 48106

18 Bedford Row, London WC1R 4EJ, England
THE LEAST PRIME IDEAL WITH PRESCRIBED
DECOMPOSITION BEHAVIOUR

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in The Graduate
School of The Ohio State University

BY
Alfred R. Weiss

* * * * * * *
The Ohio State University
1980

Reading Committee:
M. Madan
P. Ponomarev
H. Zassenhaus

Approved By

[Signature]
Adviser
Department of Mathematics
ACKNOWLEDGMENTS

My thanks must go first to my wife for her patience and understanding. I would also like to thank the faculty members who offered encouragement and especially my advisor, Professor Hans Zassenhaus, for his insight and friendly helpfulness.
VITA

November 20, 1950.............. Born - Saskatoon, Saskatchewan

1971......................... B.Sc., University of Alberta, Edmonton, Alberta

1971-1974..................... Teaching Associate, Department of Mathematics
Ohio State University, Columbus, Ohio

1973......................... M.S., Ohio State University, Columbus, Ohio

1974-1975..................... Teaching Associate, Department of Mathematics
California Institute of Technology, Pasadena, California

1975-1978..................... Teaching Associate, Department of Mathematics
Ohio State University

PUBLICATIONS


FIELDS OF STUDY

Major Field: Mathematics

Studies in Number Theory. Professor R. Gold, Professor M. Madan,
Professor A. Woods, and Professor H. Zassenhaus

Studies in Algebra. Professor J. Ferrar and Professor H. Zassenhaus

Studies in Analysis. Professor B. Baishanski
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgments</td>
<td>11</td>
</tr>
<tr>
<td>Vita</td>
<td>iii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter I - Basic Estimates</td>
<td>11</td>
</tr>
<tr>
<td>§1 - Notation and the functional equation of Hecke</td>
<td>11</td>
</tr>
<tr>
<td>§2 - A Phragmén-Lindelöf estimate and the Hadamard factorization</td>
<td>18</td>
</tr>
<tr>
<td>§3 - The non-trivial zeros of $L(s,\chi)$</td>
<td>27</td>
</tr>
<tr>
<td>§4 - The exceptional zero</td>
<td>42</td>
</tr>
<tr>
<td>Chapter II - A special arithmetic sum</td>
<td>48</td>
</tr>
<tr>
<td>§1 - A lower bound</td>
<td>48</td>
</tr>
<tr>
<td>§2 - A upper bound</td>
<td>55</td>
</tr>
<tr>
<td>§3 - An analogous sum for prime ideals</td>
<td>62</td>
</tr>
<tr>
<td>Chapter III - Mean values of Dirichlet polynomials</td>
<td>67</td>
</tr>
<tr>
<td>§1 - Averaging functions and their properties</td>
<td>67</td>
</tr>
<tr>
<td>§2 - An inequality of Polya-Vinogradov type, for ideals</td>
<td>75</td>
</tr>
<tr>
<td>§3 - Application of Selberg's method</td>
<td>87</td>
</tr>
<tr>
<td>§4 - Conclusion</td>
<td>94</td>
</tr>
<tr>
<td>Chapter IV - The zero-density theorem</td>
<td>101</td>
</tr>
<tr>
<td>§1 - Detecting non-exceptional zeros</td>
<td>101</td>
</tr>
<tr>
<td>§2 - The main theorem</td>
<td>121</td>
</tr>
<tr>
<td>Chapter V - Arithmetical applications</td>
<td>136</td>
</tr>
<tr>
<td>§1 - An explicit formula</td>
<td>136</td>
</tr>
<tr>
<td>§2 - Prime ideals in short segments of arithmetic progressions</td>
<td>148</td>
</tr>
<tr>
<td>§3 - Prime ideals in Tchebotarev classes</td>
<td>165</td>
</tr>
<tr>
<td>§4 - Examples</td>
<td>174</td>
</tr>
<tr>
<td>List of References</td>
<td>177</td>
</tr>
</tbody>
</table>
There are two problems with which we shall be concerned: briefly, the problem of the least prime in a (generalized) arithmetic progression for an arbitrary algebraic number field $K$, and the problem of the least prime having given decomposition behavior (i.e. belonging to a specified Tchebotarev class) in a finite galois extension $L/F$ of algebraic number fields. To fix ideas we formulate each of these problems more precisely.

Let $K$ be an algebraic number field of finite degree $n$ over $\mathbb{Q}$ and with absolute discriminant $d_K$. For each integral divisor $m$ of $K$ we let $I(m)$ denote the group of fractional ideals of $K$ prime to $m$, and let $P_m$ denote the ray modulo $m$ i.e. the subgroup of $I(m)$ consisting of principal ideals $(\alpha)$ which admit a generator satisfying: $\alpha \equiv 1 \mod m$ and $\alpha$ is positive at every real archimedean "prime" of $K$. Then the first problem is

**PROBLEM AP**: find a bound $M'(m)$ depending on the arithmetic of $K$ and on $m$ so that in every coset of $P_m$ in $I(m)$ there is a prime ideal $p$ satisfying $Np \leq M'(m)$.

Here $N$ is the absolute norm on $K$; the focal point of this work is the proof of

**THEOREM AP**: $M'(m) = C_1 \max\left\{ (d_K N(m))^{C_2/n}, C_3 n\right\}$ solves problem AP for suitable constants $C_1$, $C_2$, $C_3$ which are effectively computable and
independent of the field $K$.

Actually, in theorem 5, we will prove somewhat more than this, in two respects. The first improvement is that we shall, in fact, prove that for every $x \geq \frac{1}{2} M_K(m)$ (i.e. that given in theorem AP) and in each coset of $P_m$ in $I(m)$ there is a prime ideal $p$ so that

$$x < Np < x + x^{1-C_4/n}$$

where $C_4$ is, again, effectively computable and independent of $K$. Indeed we will even give a lower bound of this type for the number of primes satisfying these conditions, by using the effective replacement for Siegel's theorem which has been given by Stark [27].

The second improvement is in the formulation we will give to theorem AP (of which the given version is a special case). The motivation behind the formulation is that it be suited to the solution of our second basic problem which we will first formulate.

Let $L/F$ be a finite galois extension of algebraic number fields with galois group $G$. If $p$ is a prime ideal of $F$ which is unramified in $L$ then the Frobenius automorphisms of the prime ideals $P$ of $L$ which lie above $p$ fill a single conjugacy class of $G$ which we denote $(p, L/K)$. The second problem is

**PROBLEM TC**: Given a conjugacy class $C$ of $G$ find a bound $M'_{L/F}(C)$ depending on the arithmetic of the extension $L/F$ (and possibly $C$) so that there exists a prime ideal $p$ of $F$ satisfying

$$(p, L/F) = C \quad \text{and} \quad Np \leq M'_{L/F}(C).$$
In the special case when $G$ is abelian, hence $C$ is a single element of $G$, the problem $TC$ almost reduces to the problem $AP$, via class field theory. We say "almost" because the formulation of theorem $AP$ amounts to considering the conductor of an abelian extension as the only invariant of the extension, again via class field theory.

Our second improvement, then, is an attempt to solve problem $AP$ in such a way that the invariants involved transfer to the problem $TC$. The idea is to focus on the Artin conductors of the irreducible representations of $G$ as the arithmetic invariants of $L/F$ on which to build $M'_{L/F}(C)$.

The means of reducing problem $TC$ to problem $AP$ is a simple argument of Deuring for the Tchebotareff density theorem which we will repeat in §3: roughly, one takes any abelian subgroup $A$ of $G$ which meets $C$, forms the fixed field $K$ of $A$, and interprets the abelian extension $L/K$ in terms of class field theory. Then the problem $TC$ can be reduced to problem $AP$ in this field $K$, and these two problems are then essentially equivalent, except for one thing: in the resulting problem $AP$ we "forget" the way in which $K$ is related to $L/F$. However, in the absence of any real theory of non-abelian extensions, this seems unavoidable at the present time.

What we do get for problem $TC$, in theorem 6, is a bound depending on Artin conductors of representations of $G$ induced from $A$ and involving effective constants as in theorem $AP$. Such a bound is obtained for any abelian subgroup $A$ of $G$ which meets $C$, and roughly speaking, the bound improves with the size of such an $A$ (see §3).
With this reduction step out of the way we now describe our approach to theorem AP and begin by pointing out that theorem AP for \( K = \mathbb{Q} \) was proved by Yu.V. Linnik [19]. Thus our main interest is focused (as it must be for the application to problem TC) on a general result for every field \( K \) which keeps explicit track of the dependence on \( K \).

Such an approach to problem AP was taken by Fogels [8] although he did not attempt to control the influence of the degree \( n \) of \( K \) over \( \mathbb{Q} \). Building on the work of Fogels, but returning to \( K = \mathbb{Q} \), Gallagher [10] introduced the integration technique which we use in IV§1 and made some simplifications thereby giving a greatly simplified treatment of Linnik's theorem. Bombieri [2] then observed that the treatment of the exceptional zero (see I§3) could be incorporated in the same argument, instead of requiring special handling as it always had before; Bombieri also considers only the case \( K = \mathbb{Q} \). However, both Bombieri and Gallagher emphasize the large sieve and show how, in its formulation by Bombieri-Davenport [3] with improvements by Gallagher [9], this tool applies to Linnik's theorem. Although the large sieve has been generalized to algebraic number fields, in particular by Huxley [4,15], Wilson [30], and Schaal [25], in such a way that, with some additional adjustments, it can be applied to problem AP (for general \( K \)) it does not yet seem possible to make this application in a manner compatible with the requirement of getting effective bounds in terms of arithmetic invariants that are recognizable.
Because of these problems we return to the method of Fogels who uses the Selberg sieve in place of the large sieve. However in Fogels' treatment there is no possibility of getting a result better than the $n^{th}$ power of our theorem $AP$ (and perhaps worse; this is not clear since no attempt is made at controlling $n$, as mentioned above). The reason for this is that the application of the Selberg sieve is based on estimates for the distribution of integral ideals in cosets of $I(m)/P_m$; oversimplifying somewhat we point out that the best results for the number of integral ideals of $K$ which have norm $\leq x$ are of the form

$$\kappa(K)x + O_K(x^{1-2/n+1})$$  \hspace{1cm} (2)

where the constant implied by $O_K$ depends on $K$ (this particular result goes back to Landau [17]), and, in the interest of $K$-uniform results, Fogels improves this to

$$\kappa(K)x + O_n(d_K^{2/3}x^{1-2/n+3})$$  \hspace{1cm} (3)

where $O_n$ now depends only on $n$. Now, since by the Brauer-Siegel theorem the residue $\kappa(K)$ of $\zeta_K(s)$ at 1 can be expected to be "close to 1" (the precise meaning of this will cause many of the later problems), it is clear that the error term of (3) is the dominant term unless $x$ is significantly larger than $d_K^n$. What is worse is that this is unavoidable (without assuming the Generalized Riemann Hypothesis, which is pointless here because on that assumption it is possible to get a much better bound for $M'_L/F(C)$ of problem $TC$ directly; see Lagarias-Odzko [16]) in the sense that if we want an
upper bound for the number of prime ideals in a coset of \( I(m)/P_m \) by the Selberg sieve then we are forced to use (3) or some improvement of it, a result we do not presently know. Finally this difficulty with (3) is not confined to the application of the Selberg sieve but appears every time we need to estimate arithmetic sums over integral ideals.

However, we are primarily interested in prime ideals and although no improvement will be made on (3) it turns out that no improvement will be needed in that we can avoid the issue altogether. The underlying idea is that of averaging (or, perhaps more precisely, smoothing) the arithmetic sums we must consider i.e. replacing a sum \( \sum_{d} b(d) \) by a sum \( \sum_{d} b(d)f(N(d)) \) where \( f \) is a "nicely behaved" function. Roughly speaking "nicely behaved" will amount to the Mellin transform of \( f \) being analytic and bounded in a suitable vertical strip in the complex plane and vanishing quickly enough (depending on the degree \( n \) of \( K \) over \( \mathbb{Q} \)) at \( \infty \) inside the strip. Of course this idea in itself is of little use: the point is that such averaged sums behave better than (3) in the sense that the error term is the smaller term whenever \( x \geq d_K \) for an effective constant \( C_5 \) independent of \( K \) (warning: this is only an analogy for comparison with (3), just as (3) itself is a gross simplification of the sums we must consider, except in chapter II) and that by suitable (strictly ad hoc) devices we can reduce to the consideration of such "averaged" sums.

In particular for some of these sums we only need upper (or lower) bounds: then the technique of Landau [17], which he used to prove (2), is a suitable device, and this case occupies chapter II.
Again, the integration technique of Gallagher mentioned above leads to the consideration of certain mean values of Dirichlet "polynomials" (i.e. truncated Dirichlet series, at least heuristically) and a suitable extension of Gallagher's basic lemma for handling such mean values allows us to replace his sums over "short intervals" by averaged sums over similar "short intervals". As Gallagher points out, Fogels' proof can be simplified by combining his method with the Brun-Titchmarsh inequality: in the same way our extension of his method combines with an "averaged form of the Brun-Titchmarsh inequality" (see lemma 9) which we obtain, via the Selberg sieve, from an "averaged form of the Polya-Vinogradov inequality" (which is inspired by appendix II of Montgomery [20]). Carrying out this program is the object of chapter III. Finally, this theme reoccurs one last time in V§1§2 when we prove theorem AP (in the stronger form mentioned above) by using an "averaged explicit formula" in place of the more usual explicit formula of prime ideal theory: this again is based on an idea of Fogels who uses a somewhat different "averaged explicit formula" with the purpose of proving the existence of prime ideals in the "short interval" (1), although as usual his constants depend on $n$ in an undetermined way. For our purposes the "averaged explicit formula" also allows an important technical maneuver, the explanation of which is deferred to the paragraph preceding the corollary to lemma 13.

Although "averaging" is the main idea in the whole proof it also contributes what must be considered the major flaw in theorem AP: namely the occurrence of $n^n$ in our bound $M_K(m)$. Thus while $d_KN(m) \geq d_K > c_6^n$ for some $c_6 > 1$ (provided $K \neq \emptyset$) by the
Minkowski lower bound for discriminants there are indeed fields $K$ for which $d_K$ is significantly smaller than $n^n$, notably the fields in infinite Hilbert class field towers (which exist by Golod-Safaravic). On the other hand the occurrence of $n^n$ seems to be a natural consequence of our averaging processes, and, remarkably, the biggest flaw in the explicit Siegel theorem of Stark [27] is a similar occurrence of $n^n$ although it appears there for an entirely different reason (for which, see the paragraph following theorem 1').

Similarly the reliance on the Selberg sieve also causes an attendant difficulty, via the need to estimate the number of integral ideals. This is discussed in more detail in the paragraph of (17), (18) of V§2. Yet some sort of sieving argument has been essential in every attempt to solve problem AP to date.

The methods of chapters I and IV are almost purely analytical, except in the sense that they exploit the arithmetical information which is encoded in the functional equation of Hecke for $L$-series. In particular in chapter I we prove $K$-uniform versions of the basic facts about $L$-series e.g. a Phragmen-Lindelöf bound for $L(s,\chi)$ in the critical strip (due to Fogels [6] who obtains (3) from it), the "usual" zero-free region (appropriately normalized), and various zero density results (again appropriately normalized). Finally in I§4 we sketch Stark's results on the explicit Siegel theorem, which ultimately leads to the effective lower bound for the number of primes $p$ satisfying (1). Also the significance of the relation (14) of I§2, on which so much of the $K$-uniformity hinges, was first realized by Stark [27].
Finally we come to chapter IV which culminates in theorem 4 which we must regard as the main result: indeed everything preceding theorem 4 (except i54) is used in the proof. Theorem 4 is a zero density theorem which says that the number of zeros of $L(s,\chi)$ (with the exception of a single "exceptional" zero) near the line $\text{Re}(s) = 1$ is "small". The results of IV§1 are, with the exception of the treatment of the $\Phi_X$ term, K-uniform versions of the work of Fogels-Gallagher-Bombieri. Because the proof of theorem 4 uses an integration technique we are able to handle the $\Phi_X$ term (which comes from the pole of $\zeta_K(s)$ at $s = 1$) in a manner compatible with the other terms, which fact is extremely important to preserving K-uniformity since otherwise the pole of $\zeta_K(s)$ at $s = 1$ seems to "hide" the "exceptional" zero, when it exists.

The importance of theorem 4 stems from the fact that it supplies the information needed for the analysis of the "average explicit formula" mentioned, and even though the arguments of chapter IV are mainly analytical it must be emphasized again that its proof depends at a critical stage on a sieving method, and that the applicability of the Selberg sieve, in particular, depends on the way the prime ideals are related to the integral ideals, and on our choice of averaging process.

The reason for pointing this out again is that it may seem, in view of the majority of the arguments being analytical, that it may be possible to attack problem TC by using the Artin L-series directly (and it seems that ultimately this will have to be the way), assuming that they are entire. While there are some steps that can be taken
similarly to the Hecke L-series case we consider (e.g. the identity 5) of $I_n^3$ is a suitable generalization of the usual trigonometric identity used in deriving the classical zero free regions) there are far more steps which do not work at all e.g.

i) How are real-valued characters of representations to be compared in the context of lemma 2?

ii) How is the Selberg sieve argument to be simulated? More specifically is there a set containing the prime ideals which in some sense simulates the set of integral ideals for the purpose of applying the Selberg sieve?

iii) Can the Artin L-series of an irreducible representation of degree $d > 1$ be factored into $d$ entire functions thus making the analytical "weight" of the factors comparable to the "weight" of Hecke L-series?

The question (iii) has received some (so far inconclusive) attention from Vinogradov [29], Goldfeld [11]. However in the absence of any information on such questions it appears we must, at this time, be content with reduction to the abelian case.
CHAPTER I

BASIC ESTIMATES

The purpose of this chapter is to develop analytical estimates of various kinds for Hecke $L$-series defined over a fixed algebraic number field. Nevertheless the constants involved in these estimates are always effectively computable and independent of the field in question. The methods employed are mainly classical with various modifications necessitated by the requirement that the estimates be independent of the field, which is critical for the later applications.

§1 Notation and the functional equation of Hecke

Let $K$ be an algebraic number field of finite degree $n = n_K$ over $\mathbb{Q}$ and absolute discriminant $d_K = d_{K/\mathbb{Q}}$. Let $r_1$, respectively $r_2$, denote the number of real, respectively complex, completions of $K$, hence $r_1 + 2r_2 = n$; and let $N$ denote the absolute norm: thus if $a$ is a fractional ideal of $K$ then $N(a)$ is the positive generator of the principal fractional $\mathbb{Q}$-ideal $N_{K/\mathbb{Q}}(a)$. The field $K$ will remain fixed unless an explicit statement is made to the contrary.

If $m$ is an integral ideal of $K$ let $K(m)$ denote the (multiplicative) group of elements $\alpha$ of $K$ so that the prime ideals which appear in the principal fractional $K$-ideal $(\alpha)$ do not divide $m$. For $\alpha \in K(m)$ the multiplicative congruence

$$\alpha \equiv 1 \mod \mathfrak{m}$$
will mean that, first, \( \alpha \) is totally positive (i.e. positive in every real completion of \( K \)) and, second, that for every prime ideal \( p \) which divides \( m \) with some multiplicity \( m(p) > 0 \) the element \( \alpha \) of the local ring \( \mathcal{O}_p \) at \( p \) represents the unit element of the quotient ring \( \mathcal{O}_p / (p0p)^{m(p)} \). We put \( K_m = \{ \alpha \in K(m): \alpha \equiv 1 \mod x m \} \).

Similarly, \( I(m) \) will denote the group of all fractional ideals whose prime factors do not divide \( m \) and \( \Gamma_m \) is the subgroup of \( I(m) \) consisting of those principal fractional ideals \( (\alpha) \) which have a generator \( \alpha \) in \( K_m \). The quotient group \( I(m)/\Gamma_m \) is known as the group of \( m \)-ideal classes (of \( K \)). It should be observed that we have avoided the finer distinction which arises when in the definition of \( K_m \) we insist that \( \alpha \) be positive only at some of real completions with no condition on the others: thus when \( m = 0 \), the unit ideal, the class group \( I(\mathcal{O})/\Gamma_0 \) is the group of ideal classes in the strict sense. The reason for doing this is that our later estimates do not become significantly sharper on preserving the distinction.

In the above context we refer to \( m \) as a modulus to emphasize that it is only the partial ordering on the moduli induced by the divisibility relation (i.e. \( m \geq n \) if and only if \( m \) divides \( n \)) that concerns us.

By a character \( \chi \mod m \) we understand a homomorphism \( \chi \) of \( I(m)/\Gamma_m \) into the complex numbers of absolute value 1. If the modulus \( n \) is larger than \( m \) (i.e. \( m \) divides \( n \)) then the inclusion \( I(n) \to I(m) \) induces a surjective homomorphism

\[ I(n)/\Gamma_n \to I(m)/\Gamma_m \]
which when composed with a character $χ \mod m$ yields a character $ψ \mod n$: we say $χ$ induces $ψ$. In the case when $χ \mod m$ is induced by no character with strictly smaller modulus then $χ$ is called a primitive character $\mod m$. In general, every character $χ \mod m$ is induced by a unique primitive character $ψ \mod \delta$ where $\delta$ dividing $m$ is the smallest modulus of any character inducing $χ$: $\delta$ is called the conductor of $χ$ (and of $ψ$) and denoted $\delta_χ$. For a particularly clear account of this see Heilbronn [12].

We next summarize the main qualitative properties of the $L$-series while fixing notation. For proofs we refer to Lang [18] by chapter and section (since the relevant facts appear scattered throughout his book).

For a character $χ \mod m$ one defines

$$L(s,χ) = \sum_\mathfrak{a} χ(a)N(a)^{-s}, \quad \text{Re}(s) > 1$$

where the sum is over all integral ideals in $I(m)$, or, equivalently, over all integral ideals on adopting the usual convention that $χ(\mathfrak{a}) = 0$ if $\mathfrak{a}$ is not prime to $m$. The series converges absolutely for $\text{Re}(s) > 1$, and also uniformly for $\text{Re}(s) \geq 1 + \delta$ when $\delta > 0$ is fixed.

Moreover there is an Euler product representation

$$L(s,χ) = \prod_{\mathfrak{p} | \delta} (1-χ(\mathfrak{p})N(\mathfrak{p})^{-s})^{-1}, \quad \text{Re}(s) > 1$$

which again converges absolutely for $\text{Re}(s) > 1$. 
From this it is immediate that if \( \chi \mod m \) is induced by the primitive character \( \chi^* \mod m \), then we have

\[
L(s, \chi) = \prod_{p \nmid m, p \neq \delta} \left( 1 - \chi^*(p)\mathbb{N}(p)^{-s} \right) L(s, \chi^*)
\]

(3)

for \( \Re(s) > 1 \); this allows us to reduce most considerations to primitive characters.

Thus let \( \chi \) be a primitive character \( \mod m \) and put

\[
\delta_{\chi} = \begin{cases} 1 & \text{if } \chi = 1 \\ 0 & \text{if not} \end{cases}
\]

and \( \mu_{\chi} \) the number of ramified real "primes" of \( \chi \) (hence \( 0 \leq \mu_{\chi} \leq r_1 \) which is all we use). Further putting

\[
d_{\chi} = d_{K}\mathbb{N}(\delta_{\chi}), \quad A_{\chi} = 2^{-r_2} n^{-1/2} \delta_{\chi}
\]

\[
\Gamma_{\chi}(s) = \Gamma\left(\frac{s}{2}\right) \Gamma\left(\frac{s}{2} + \frac{1}{2}\right) \chi(s) \Gamma(s) \Gamma(s)
\]

(4)

we can state the fundamental result we need: \( s \chi(s-1) \chi\Lambda(s, \chi) \) has an analytic continuation to the complex plane where it is an entire function of order 1 and satisfies the functional equation

\[
\Lambda(1-s, \chi^*) = W(\chi)\Lambda(s, \chi)
\]

(5)

where \( W(\chi) \) is a constant of absolute value 1 (which will not otherwise concern us). This fundamental result, due to Hecke, can be found in chapter XIV and the order 1 result in chapter XVIII.1 of Lang.
Finally to complete this we need to know that
\( (s-1) \chi L(s, \chi) \) takes a non-zero value at \( s = 1 \). \( (6) \)

This amounts to \( L(1, \chi) \neq 0 \) for \( \chi \neq 1 \) and to \( L(s, \chi) \) having a simple pole at \( s = 1 \) when \( \chi = 1 \), and can be found in chapter VII or in chapter XV§4 of Lang. Using this we can describe the placement of the zeros of \( L(s, \chi) \) because by
\[
\Gamma(z)^{-1} = z e^{\gamma z} \prod_{m=1}^{\infty} \left( 1 + \frac{z}{m} \right) e^{-z/m}
\](7)

the \( \Gamma \)-function has no zeros and has simple poles at the non-positive integers, hence \( \Gamma_\chi(s) \) has no zeros and all of its poles are at (some of the) nonpositive integers (with some multiplicity).

Now by the absolute convergence of the Euler product (2) \( L(s, \chi) \) has no zeros (or poles) in \( \text{Re}(s) > 1 \), hence by (4) and the above information on \( \Gamma_\chi(s) \) we see that \( \Lambda(s, \chi) \) has no zeros in \( \text{Re}(s) > 1 \), from which the functional equation (5) (applied to \( \chi \)) shows that \( \Lambda(s, \chi) \) has no zeros in \( \text{Re}(s) < 0 \). Also by (4) and (6) together with \( \Gamma_\chi(1) \neq 0 \) we see that \( \delta s \chi \Lambda(s, \chi) \) takes a non-zero value at \( s = 1 \), hence by (5) \( s \chi \Lambda(s, \chi) \) takes a non-zero value at \( s = 0 \) so that

all zeros of \( \Lambda(s, \chi) \) satisfy
\[
0 \leq \text{Re}(s) \leq 1 \text{ and } s \neq 0,1.
\] \( (8) \)

Moreover since \( \Gamma_\chi \) has no zeros, (4) and (5) show that \( \delta s \chi (s-1) \chi L(s, \chi) \) is entire. But this says
$L(s, \chi)$ is entire when $\chi \neq 1$; and when $\chi = 1$, $L(s, \chi)$ is meromorphic with only a single simple pole at $s = 1$, (9)

for in the outstanding case $\delta_\chi = 1$, $s = 0$ we already know

$s^\chi \Lambda(s, \chi)$ takes a non-zero value at $s = 0$ hence so does $s^\chi \Gamma_\chi (s)L(s, \chi)$ by (4) when the conclusion follows since $\chi = 1$ implies $\mu_\chi = 0$ hence $\Gamma_\chi$ has a pole at $s = 0$.

Thus all the zeros of $L(s, \chi)$ outside the region defined by (8) must arise from poles of $\Gamma_\chi(s)$ i.e. we have seen that $s^\chi (s-1)\Lambda(s, \chi)$ is never zero outside the region (8) hence by (9) and (4) it follows that $s^\chi \Gamma_\chi (s)L(s, \chi)$ is never zero outside the region defined by (8) and from the known poles of $\Gamma_\chi(s)$ we can deduce that all such zeros of $L(s, \chi)$ are at (some of the) non-positive integers, and even deduce their multiplicities from the non-vanishing of $s^\chi \Gamma_\chi (s)L(s, \chi)$ . These (known) zeros outside the region (8) are called trivial zeros and will not concern us further.

On the other hand zeros of $L(s, \chi)$ in the region (8) i.e. satisfying $0 \leq \text{Re}(s) \leq 1$, $s \neq 0, 1$, are then non-trivial zeros. By the equation (4) and the fact that $\Gamma_\chi(s)$ has no poles in this region we can summarize by saying:

The non-trivial zeros of $L(s, \chi)$ correspond (including multiplicities) exactly to the zeros of $\Lambda(s, \chi)$ : they form a "set" (with multiplicities) we denote $\mathbb{L}(\chi)$ . All other zeros of $L(s, \chi)$ are at (some of the) non-positive integers.
Of course when \( \chi = 1 \) \( L(s,\chi) \) is nothing but the Dedekind zeta function \( \zeta_K(s) \) so (9) assures us that we have a Laurent expansion

\[
\zeta_K(s) = \frac{\kappa(K)}{s-1} + \kappa_0(K) + \kappa_1(K)(s-1) + \ldots
\]

with \( \kappa(K) \neq 0 \), which must converge in the entire (finite) complex plane since \( \zeta_K(s) \) has no other singularities.

Returning, briefly, to imprimitive characters \( \chi \mod m \) we note that equation (3) can now be considered as giving an analytic continuation of \( L(s,\chi) \).

Finally it is important to clarify our use of implied and unspecified constants in what follows: we use the (equivalent) notations

\[
f(x) \ll g(x) \text{ , } f(x) = O(g(x))
\]

over some specified domain of \( x \) to assert the existence of an effectively computable (in the naive sense that numerical values could be assigned) constant \( C \) which is independent of \( K \) so that \( |f(x)| \leq Cg(x) \) over that domain of \( x \). Certain constants of this type will be singled out and denoted \( c_1, c_2, \ldots \): once such a \( c_m \) is specified it is considered fixed from then on, and constants which appear later may depend on that value (also: definitions).

This usage is crucial to all results which follow and will be in effect unless the contrary is explicitly stated, a situation which will occur when we prove technical results e.g. the convergence of some integral. Thus we signal that we are using \( \ll \) or \( \mathcal{O} \) outside of the above framework by the catchword "qualitatively."
52 A Phragmén-Lindelöf estimate and the Hadamard factorization

Throughout this § we fix a primitive character $\chi \mod \ell$. Also we shall use (from now on) the traditional notation $s = \sigma + it$ for a complex variable $s$. Our first task will be to give an upper bound for $L(s, \chi)$ in the strip $0 \leq \sigma \leq 1$ which we base on Stirling's formula

$$
\log \Gamma(s) = (s - \frac{1}{2}) \log s - s + \frac{1}{2} \log 2\pi + O\left(\frac{1}{|s|}\right)
$$

which holds uniformly in the region $|\arg s| \leq \pi - \delta$, where $\delta > 0$ is fixed, excluding neighborhoods of the non-positive integers (where $\Gamma$ has poles). Here, and throughout, log and arg denote the principal branch of these functions i.e. arg takes values in $(-\pi, \pi]$.

The result in question is due to Fogels [6] in the form we need i.e. uniform in the whole doubly infinite strip, and is

**Lemma 1.** For $a \geq 1$ we have

$$
\left| \left(\frac{s-1}{s+2}\right)^{\frac{1}{2}} L(s, \chi) \right| \leq e^{O(n)} a^{\frac{1}{2} \left(1+\frac{1}{a} - \sigma\right)(1+|t|)^2} e^{\frac{n}{2} \left(1+\frac{1}{a} - \sigma\right)}
$$

uniformly in the strip $-\frac{1}{a} \leq \sigma \leq 1 + \frac{1}{a}$. Moreover the constant implied in $e^{O(n)}$ is independent of $a \geq 1$.

**Proof:** Because of the importance of this result for later estimates we give the proof in detail. First from the Stirling formula we get,

$$
|\Gamma(s)| = e^{O(1)} (1+|t|)^{\sigma - \frac{1}{2}} e^{-\frac{\pi}{2} |t|}
$$

uniformly in $1 \leq \sigma \leq 2$ (2)
which is all we will need (for the deduction of (2) from (1) see Prachar [23], p. 395).

Now by the Dirichlet series (1) of §1 we have

$$|L(s, \chi)| \leq \zeta_K(\sigma), \quad \sigma > 1$$

and by the Euler product (2) of §1 also

$$\zeta_K(\sigma) \leq \zeta_Q(\sigma)^n, \quad \sigma > 1$$

by considering the prime ideals $p$ which divide a fixed prime $p$ of $\mathbb{Q}$. But by partial summation (given in Lang [18], p. 157)

$$1 \leq (\sigma-1)\zeta_Q(\sigma) \leq \sigma, \quad \sigma > 1$$

so putting these together and using $1 + a \leq e^{O(1)}a$ (by $a \geq 1$) we can conclude that

$$|L(s, \chi)| \leq e^{O(n)}a^n \text{ on the line } \sigma = 1 + \frac{1}{a}. \quad (3)$$

Next we need such an estimate on the line $\sigma = -\frac{1}{a}$ for which we use the functional equation (5) of §1 in the form

$$L(s, \chi) = \frac{\zeta(1-s)}{\zeta(s)} L(1-s, \chi).$$

Since $|\zeta(s)| = 1$ and $A_{\chi}^{1-2\sigma} = e^{O(n)}\frac{1}{\sqrt{\chi}} + \frac{1}{a}$ on the line $\sigma = -\frac{1}{a}$ and since the equation (3) applies to $L(1-s, \chi)$ we are left to estimate $\Gamma$-factors. Since the estimate (2) is somewhat awkward to apply directly it is perhaps easier to use the relations
\[
\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}, \quad \Gamma(s)\Gamma(s+\frac{1}{2}) = 2^{1-2s} \pi^{1/2} \Gamma(2s)
\]
to write

\[
\frac{\Gamma(1-s)}{\Gamma(s)} = \pi^{-n/2} r_1^s (\sin \frac{\pi s}{2})^{1-\mu} \chi(\sin \frac{\pi}{2}(1+s))^{\mu} \chi(\sin \pi s)^{r_2} \Gamma(1-s)^n.
\]

Now since \( \sigma = \frac{-1}{\alpha} \) puts \( 1 - s \) in the region in which (2) is valid
and since \( |\sin s| = \left| \frac{e^{is} - e^{-is}}{2i} \right| \leq e |t| \) we get immediately

\[
\left| \frac{\Gamma(1-s)}{\Gamma(s)} \right| \leq e^{O(n)} e^{\frac{n}{2}|t|} \left( (1+|t|)^{\frac{1}{2} + \frac{1}{\alpha}} e^{-\frac{n}{2} |t|} + O(1) \right)^n.
\]

for \( \sigma = \frac{-1}{\alpha} \),

so putting the pieces together yields

\[
|L(s, x)| \leq e^{O(n)} \frac{1}{\alpha} (1+|t|)^{n \left( \frac{1}{2} + \frac{1}{a} \right)} a^n \text{ on } \sigma = \frac{-1}{\alpha}.
\]

We begin the construction of the auxiliary function for the application of Phragmén-Lindelöf results by defining

\[
g(s) = \frac{\Gamma(2-s)}{\Gamma(1+s)} \quad \text{on } 0 \leq \sigma \leq \frac{1}{2}
\]

where this function is holomorphic and non-vanishing. Moreover the estimate (2) applies to both \( \Gamma \)-factors and yields

\[
|g(s)| = e^{O(1)} (1+|t|)^{1-2\sigma}, \quad 0 \leq \sigma \leq \frac{1}{2}.
\]

The auxiliary function is now obtained by putting
\[ G(s) = \exp\left(-n\left(\frac{1}{2} + \frac{1}{a}\right) \log \left(\frac{s + \frac{1}{a}}{2\left(1 + \frac{1}{a}\right)}\right)\right), \quad -\frac{1}{a} \leq \sigma \leq 1 + \frac{1}{a} \]

and from the estimate for \(|g(s)|\) we get

\[ G(s) = e^{\Theta(n)} \left(1 + \frac{a}{2(a+2)} |t|\right)^{-n\left(1 - \frac{1}{a} - \sigma\right)}, \quad -\frac{1}{a} \leq \sigma \leq 1 + \frac{1}{a}. \quad (6) \]

Putting \(f(s) = (s-1)^{-1} L(s, \chi) G(s), \quad -\frac{1}{a} \leq \sigma \leq 1 + \frac{1}{a}\), observing that

\[ \frac{1}{6} \leq \frac{a}{2(a+2)} \leq \frac{1}{2} \quad \text{and that} \quad \left|\frac{s-1}{s+2}\right| \quad \text{is bounded independently of} \quad a \quad \text{on} \]

\(\sigma = -\frac{1}{a} \quad \text{and} \quad \sigma = 1 + \frac{1}{a} \quad \text{(by} \quad a \geq 1) \quad \text{we have, by (3) and (6)} \]

\[ |f(s)| \leq e^{\Theta(n)} a^n \quad \text{on the line} \quad \sigma = 1 + \frac{1}{a} \quad (7) \]

and, by (5) and (6),

\[ |f(s)| \leq e^{\Theta(n)} a^{|\gamma| + \frac{1}{a}} \left(1 + |t|\right)^{-n\left(1 - \frac{1}{a} - \sigma\right)} \quad (8) \]

\[ = e^{\Theta(n)} a^{|\gamma| + \frac{1}{a}} \quad \text{on the line} \quad \sigma = -\frac{1}{a}. \]

But \(f(s)\) is analytic in the strip \(-\frac{1}{a} \leq \sigma \leq 1 + \frac{1}{a}\) (since the factor \((s-1)^{-1}\) cancels the only possible pole), is bounded on the sides by (7), (8) and finally \(f(s)\) is built from \(\Lambda(s, \chi)\) and \(\frac{1}{\Gamma(s)}\)

which are entire functions of order 1 so \(f(s)\) satisfies the weak growth condition \(|f(s)| \propto |t|^{1+\varepsilon}\) as \(|t| \rightarrow \infty\) (for any \(\varepsilon > 0\)),

hence by the Phragmén-Lindelöf theorem \(f(s)\) is bounded in the strip \(-\frac{1}{a} \leq \sigma \leq 1 + \frac{1}{a}\) (by Lang [18], p.262, for example). This is so far only a qualitative bound but allows the application of the "first
convexity theorem" (Lang, p. 263) from which (7), (8) imply

\[ |f(s)|^{1+2/a} \leq \left( e^{O(n)} a n^{2} \chi + \frac{1+1}{a} \right)^{1+1/a} - \sigma \left( e^{O(n)} a n^{\sigma+1} \right) \]

throughout the strip \(-\frac{1}{a} \leq \sigma \leq 1 + \frac{1}{a}\). Simplifying this gives

\[ |f(s)| \leq e^{O(n)} a n^{2} \chi \]

and the lemma follows by the definition of \(f(s)\) and the estimate (6).

Here we have used at a critical stage the fact that \(\delta_{s} \chi(s-1) \chi_{A}(s,\chi)\) has order 1; however this same fact also gives us the Hadamard factorization

\[ \delta_{s} \chi(s-1) \chi_{A}(s,\chi) = a_{\chi} e^{\chi_{1} - \sum_{\rho \in \mathcal{L}(\chi)} (1-\frac{s}{\rho}) e^{s/\rho}} \]  

where \(\sum_{\rho \in \mathcal{L}(\chi)} \frac{1}{|\rho|} \leq 1 + \varepsilon\) converges for every \(\varepsilon > 0\), a fact that is proved very efficiently in chapter 11 of Davenport [4]. The constant \(a_{\chi}\) can be evaluated by putting \(s = 0\) and using the functional equation, however our interest lies in \(b_{\chi}\). To get \(b_{\chi}\) we need

\[ \Lambda(s,\chi) = \Lambda(s,\overline{\chi}) \]  

Since \(\chi,\overline{\chi}\) have the same kernel we have \(\Lambda_{\chi} = \Lambda_{\overline{\chi}}, \Gamma_{\chi} = \Gamma_{\overline{\chi}}\) immediately, hence it suffices to see that \(L(s,\chi) = L(s,\chi)\) which is clear from the Dirichlet series for real \(s > 1\) hence everywhere, by analytic continuation. From (10) we see that \(\overline{b_{\chi}} = b_{\chi}\) and by using also the functional equation (5) of §1 it follows that if \(\rho \in \mathcal{L}(\chi)\), then \(\overline{\rho}, 1 - \rho\) are in \(\mathcal{L}(\chi)\) and \(1 - \overline{\rho}\) is in \(\mathcal{L}(\chi)\), all with the same multiplicity. We can now investigate \(b_{\chi}\) (or at least its real part):
by logarithmic differentiation of (9) we have

$$\frac{A'}{A}(s,\chi) = \frac{-\delta}{s} - \frac{\delta}{s-1} - b\chi + \sum_{\rho \in \mathcal{L}(\chi)} \left( \frac{1}{s-\rho} + \frac{1}{\rho} \right).$$  \hspace{1cm} (11)

Now the logarithmic derivative of the functional equation reads

$$\frac{A'}{A}(s,\chi) = -\frac{A'}{A}(1-s,\overline{\chi})$$

and upon writing this out in terms of the above equation noting that \( \delta \) terms cancel, that \( b\frac{\delta}{X} = b\chi \) and that the zeros \( \rho \) in \( \mathcal{L}(\chi) \) can be uniquely written as \( 1 - \rho \) with \( \rho \in \mathcal{L}(\chi) \) we get

$$-b\chi + \sum_{\rho \in \mathcal{L}(\chi)} \left( \frac{1}{s-\rho} + \frac{1}{\rho} \right) = +b\chi - \sum_{\rho \in \mathcal{L}(\chi)} \left( \frac{1}{(1-s)-(1-\rho)} + \frac{1}{1-\rho} \right).$$

On simplifying, this gives

$$2\text{Re} \ b \chi = \sum_{\rho \in \mathcal{L}(\chi)} \left( \frac{1}{\rho} + \frac{1}{1-\rho} \right)$$

and recalling that whenever \( \rho \in \mathcal{L}(\chi) \) so is \( 1 - \rho \) we can group these terms together so since

$$\left( \frac{1}{\rho} + \frac{1}{1-\rho} \right) + \left( \frac{1}{1-\rho} + \frac{1}{1-(1-\rho)} \right) = 2\text{Re} \frac{1}{\rho} + 2\text{Re} \frac{1}{1-\rho}$$

we obtain finally

$$\text{Re} \ b \chi = \sum_{\rho \in \mathcal{L}(\chi)} \text{Re} \frac{1}{\rho}, \text{ hence } \text{Re} \ b \chi = \sum_{\rho \in \mathcal{L}(\chi)} \frac{\text{Re} (\rho)}{|\rho|^2} \geq 0 \hspace{1cm} (12)$$

Now entering the logarithmic derivative of the defining equation (4) of \( \Lambda(s,\chi) \) into (11) yields
\[
- \frac{L'(s, \chi)}{L(s, \chi)} - b_{\chi} + \sum_{\rho \in \mathcal{L}(\chi)} \left( \frac{1}{s-\rho} + \frac{1}{\rho} \right) = \frac{\delta}{s-1} \\
+ \log A_{\chi} + \frac{\delta}{s} + \frac{\Gamma'(s)}{\Gamma(s)} 
\]  
(13)

which on taking real parts becomes

\[
- \text{Re} \frac{L'(s, \chi)}{L(s, \chi)} + \sum_{\rho \in \mathcal{L}(\chi)} \text{Re} \frac{1}{s-\rho} = \text{Re} \frac{\delta}{s-1} + \log A_{\chi} \\
+ \frac{\delta}{s} + \text{Re} \frac{\Gamma'(s)}{\Gamma(s)} 
\]  
(14)

in view of (12). The important feature of (14) is that it eliminates \( b_{\chi} \) which by (12), depends on the non-trivial zeros of \( L(s, \chi) \), especially those close to \( s = 0 \). In case \( \chi \) is a real valued character then \( b_{\chi} \) is real, and taking real parts is not necessary to eliminate \( b_{\chi} \).

Of particular interest for later use is the case \( \chi = 1 \): we write \( b_K, \Gamma_K \) etc. to avoid possible confusion. Then in (13), we transpose the \( \frac{1}{s-1} \) term, let \( s = 1 \), and recall

\[
2b_K = \sum_{\rho \in \mathcal{L}(\chi)} \left( \frac{1}{\rho} + \frac{1}{1-\rho} \right) \text{ so obtaining the relation} \\
\lim_{s \to 1} \left( \frac{-\zeta_K'(s)}{\zeta_K(s)} - \frac{1}{s-1} \right) + b_K = \log A_K + 1 + \frac{\Gamma'(1)}{\Gamma(1)}. 
\]

But from the Laurent expansion (11) of \( \zeta_K \) we compute

\[
\frac{-\zeta_K'(s)}{\zeta_K(s)} = \frac{1}{s-1} - \frac{\zeta_0(K)}{\kappa(K)} + \ldots 
\]

which evaluates the limit above, and using
where $\gamma$ is Euler's constant (which values follow from the equation (7) of §1) we get

$$\frac{\Gamma'}{\Gamma}(1) = \gamma, \quad \frac{\Gamma'(\frac{1}{2})}{\Gamma}(\frac{1}{2}) = \gamma + 2 \log 2$$

which together with $b_K \geq 0$ give a lower bound for $\kappa_0(K)$. Of course similar relations hold for $\chi \neq 1$ but, again, we won't need them.

Since the right side of (14) is, with the aid of Stirling's formula, fairly well understood the application of (14) depends on information on the left side, of which there are two types. First, by logarithmic differentiation of the Euler product for $L(s,\chi)$ we obtain

$$-\frac{L'(s,\chi)}{L(s,\chi)} = \sum_{\alpha} \Lambda(\alpha) \chi(\alpha) a N(\alpha)^s, \quad \text{Re}(s) > 1$$

(16)

where $\Lambda = \Lambda_K$ is defined by

$$\Lambda(\alpha) = \begin{cases} \log N(p) & \text{if } \alpha \text{ is a power of the prime } p \\ 0 & \text{otherwise} \end{cases}$$

and second, since $\text{Re}(\rho) \leq 1$ for all $\rho \in \mathcal{L}(\chi)$ we have

If $\rho \in \mathcal{L}(\chi)$ and $\sigma > 1$ then $\text{Re} \frac{1}{s-\rho} > 0$

(17)

It is perhaps worthwhile to observe that the facts about non-trivial zeros $\rho$ of $L(s,\chi)$ assembled in §1 have all been utilized: namely $0 \leq \text{Re} \rho \leq 1$ figures in (12) and (17) while $\rho \neq 0$ is implicit in the Hadamard factorization (9); since the general product
has an $s^m$ factor. Thus $\rho \neq 1$ is encoded in the functional equation.

As a first application of (14) put $\chi = 1$, $s = 1 + r$, where $r > 0$ is "small" (independent of $K$). Since $\frac{\Gamma_K}{\Gamma_K K}(1) < 0$ (as follows from the values preceding (15)) we have $\frac{\Gamma_K}{\Gamma_K K}(1+r) < 0$ for $0 < r < c_0$ for appropriate $c_0$. Thus, by $\log A_K \leq \frac{1}{2} \log d_K$ and the observations (16), (17), the equation (14) reads $\frac{\Lambda(a)N(a)^{-1-r}}{a} \leq \frac{1}{r} + \frac{1}{2} \log d_K + 1$, hence certainly

$$\sum_{\lambda(a)/N(a)}^{1+r} \ll \frac{1}{r} + \log d_K \text{ for } 0 < r < c_0 \quad (18)$$

If we now put $r = 1/\log y$, where $y$ is large enough so (18) holds, and observe that $1 \leq \frac{e}{N(a)^r}$ whenever $N(a) \leq y$, then dropping the other terms in (18) leaves

$$\sum_{Na \leq x}^{1}{\frac{\Lambda(a)}{N(a)}} \ll \log d_K y \text{ provided } y \gg 1. \quad (19)$$

This simple estimate, crude as it appears, is useful because of its uniformity in $K$. Another such estimate will turn out to be useful: let $a \geq 1$ be arbitrary, when $\sum_{Na \leq x}^{1}{\frac{1}{a}} \leq \sum_{a}^{1+1/a} = \zeta_K(1+1/a)x^{1+1/a}$ and applying the estimate (3) gives

$$\sum_{Na \leq x}^{1} \leq e^{O(n)}a^n x^{1+1/a}, \; x > 0, \; a \geq 1 \quad (20)$$

In particular, by a theorem of Minkowski (on p. 119 of Lang [18]) every (broad) ideal class of $K$ contains an integral ideal with norm

$\leq c \frac{n^{1/2}}{d_K}$ where $C > 0$ is a suitable constant. Using this in (20)
§3 The non-trivial zeros of $L(s, \chi)$

We first derive a zero free region for the $L(s, \chi)$, where $\chi$ also varies, which reduces to the region of Landau-Page when $K = \mathbb{Q}$ (see chapter 5 of Prachar [23]). The method employed relies on the identity (14) of §2 (together with various devices which are straightforward generalizations of those employed for $K = \mathbb{Q}$) which seems to have been known for a long time; however, the realization that this identity is useful in obtaining $K$-uniform estimates is due to Stark [27]. It is convenient to consider $L(s, \chi)$ as a function of two arguments $s$ and $\chi$ where $\chi$ is a primitive character.

**Lemma 2.** Let $Q \geq 1$, $T \geq 1$ satisfy $Q^T T > 1$ and put $\zeta = \log QT^n$. Then there exists an effectively computable constant $c_1 > 0$ so that the region

$$ |t| \leq T, \sigma \geq 1 - 2c_1 \zeta, d \leq Q \quad (\chi \text{ primitive}) $$

contains at most one zero $(\rho_1, \chi_1)$ of $L(s, \chi)$. If $(\rho_1, \chi_1)$ exists then

a) $\rho_1$ is a simple real zero of $L(s, \chi_1)$

b) $\chi_1$ is a real-valued character.
Proof: We begin by estimating the right hand side of the identity (14) of §2 uniformly in \( \sigma \geq 1 \) (and in \( K \), of course): first, by differentiating Stirling's formula (equation (1) of §2) we have

\[
\frac{\Gamma'(s)}{\Gamma(s)} = \log s + O\left(\frac{1}{|s|}\right) \text{ uniformly in } \sigma \geq 1/2
\]

(note that considering differentiation in terms of the Cauchy integral formula gives an error term from an error term) hence by the definition of \( \Gamma_X \) we have

\[
\text{Re} \left( \frac{\Gamma_X'(s)}{\Gamma_X(s)} \right) = \frac{n}{2} \log |s| + O(n) \text{ for } \sigma \geq 1
\]

(1)

while clearly

\[
\log A_X = \frac{1}{2} \log d_K N(\delta_X) + O(n)
\]

(2)

both from (4) of §1.

We need to state the basic inequality for (possibly) non-primitive \( L \)-series so note that if \( \psi \bmod m \) is induced by \( \psi^* \bmod \psi \psi \) then by logarithmic differentiation of (3) of §1:

\[
\frac{-L'(s, \psi)}{L(s, \psi)} = \frac{-L'(s, \psi^*)}{L(s, \psi^*)} + \sum_{\frac{p}{m}, \psi(p) \chi \neq \psi^*(p)} \log N(p)
\]

where the absolute value of the sum is, for \( \sigma \geq 1 \)

\[
\leq \sum_{\frac{p}{m}, \psi(p) \chi \neq \psi^*(p)} \frac{\log N(p)}{N(p) - 1} \leq \sum_{\frac{p}{m}, \psi(p) \neq \psi^*(p)} \frac{\log N(p)}{2} + \sum_{Np=2} \frac{\log 2}{2} \leq \frac{1}{2} \log \frac{N(m)}{N(\delta_\psi)}
\]

+ \( O(n) \) since there are \( \leq n \) prime ideals \( p \) with \( Np = 2 \). Thus we have


\[- \Re \frac{L'}{L}(s, \chi) \leq - \Re \frac{L'}{L}(s, \psi^*) + \frac{1}{2} \log \frac{N(m)}{N(\delta_\psi)} + O(n) \text{ for } \sigma \geq 1.\]  

(3)

Now combining (1), (2), (3) with (4) of §2 (and writing \( L'(\psi), \delta_\psi \) in place of \( L'(\psi^*), \delta_\psi^* \)) gives the basic inequality

\[- \Re \frac{L'}{L}(s, \chi) + \sum_{\rho \in \Sigma(\chi)} \Re \frac{1}{s - \rho} \leq \Re \frac{\delta_\psi}{s - 1} \]

\[+ \frac{1}{2} \log d^K N^{(m)} + \frac{n}{2} \log |s| + O(n)\]

valid for \( \sigma \geq 1 \). Note that we have used \( \Re \frac{\delta_\psi}{s} \ll 1 \) and, particularly, that the \( N(\delta_\psi) \) terms cancel.

Fix a zero \((\rho_1, \chi_1)\) in the region

\[|t| \leq T, \sigma \geq 1 - 2c_1 L^{-1}, d_{\chi_1} \leq Q\]

where \( \chi_1 \) is primitive, and \( c_1 > 0 \) is yet to be determined. Write \( \rho = \beta + iy \) for zeros generally: we must show that \((\rho_1, \chi_1)\) is the only zero in the region and that (a) and (b) hold (if, in fact, \((\rho_1, \chi_1)\) exists at all).

Now we have

\[2 + |z|^2 + 4 \Re(z) + \Re(z^2) = 2(1 + \Re(z))^2 \geq 0\]  

(5)

for any complex number \( z \) (this is clear on writing \( z = x + iy \)) from which

\[-3 \Re \frac{L'}{L}(s, \chi_1^0) - 4 \Re \frac{L'}{L}(s + iy_1, \chi_1^1) - \Re \frac{L'}{L}(s + 2iy_1, \chi_1^2) \geq 0\]  

(6)
for any $\sigma > 1$. Here $\chi_1^0, \chi_1^2$ are not necessarily primitive characters mod $\chi_1$; indeed the Dirichlet series (16) of §2 for $L'/L$ converges absolutely for $\sigma > 1$ so the left hand side of (6) is

$$\sum \frac{A(a)}{\alpha N(a)^\sigma} \left( 3\chi_1^0(a) + 4 \Re \frac{\chi_1^1(a)}{i\gamma_1} + \Re \frac{\chi_1^2(a)}{2i\gamma_1} \right)$$

which has each term $\geq 0$ as is seen by applying (5) with $z = \frac{\chi_1^1(a)}{N(a)}i\gamma_1$ (since then $|z|^2 = 1$). We apply (4) in each of the 3 cases:

$$(s, \psi) = (\sigma, \chi_1^0), (\sigma + i\gamma_1, \chi_1^1), (\sigma + 2i\gamma_1, \chi_1^2)$$

and note that since $\chi_1^r$ is defined mod $\chi_1$ we have

$$\log d \frac{N(\chi_1)}{\chi_1} = \log d \frac{\chi_1}{\chi_1} \leq \log Q$$

in all 3 cases. Moreover when $1 \leq \sigma \leq 2$ we have

$$\frac{n}{2} \log |s| \leq \frac{n}{2} \log T + O(n)$$

again in all 3 cases, since $T \geq 1$. Thus in each of the 3 cases the right hand side of (4) is

$$\leq \Re \frac{\delta}{s-1} + \frac{1}{2} \log QT^\delta + O(n) \leq \Re \frac{\delta}{s-1} + \frac{1}{2} c^\delta x$$

for any constant $C > 1$ since $L \gg n$ allows us to absorb the $O(n)$ term. We fix $C > 1$ and the implied constant in $L \gg n$ is thereby determined (actually we may need to change it because $L \gg n$ is needed again). At any rate combining (4), (6) and noting that $\delta_0 \chi_1 = 1$ gives
\[
\frac{4}{\sigma - \beta_1} \leq \frac{3}{\sigma - 1} + \frac{4\delta_1 (\sigma - 1)}{x_1^2 (\sigma - 1)^2 + \gamma_1^2} + \frac{\delta_2 (\sigma - 1)}{x_1^2 (\sigma - 1)^2 + 4\gamma_1^2} + 4L, \quad 1 < \sigma \leq 2
\]  

(7)

where we have omitted all the Re \(\frac{1}{s-\rho}\) terms except
\[4 \Re \frac{1}{\sigma + iy_1 - \rho_1} = \frac{4}{\sigma - \beta_1}\]  
from \(\mathcal{L}(x_1)\), as we may by (17) of §2.

We distinguish 3 cases

i) \(x_1\) not real-valued

ii) \(x_1\) is real-valued and \(|y_1| \geq \frac{1}{7C^L}\)

iii) \(x_1\) is real-valued and \(|y_1| < \frac{1}{7C^L}\)

From (7) we deduce in cases (i), (ii) that
\[
\frac{1}{\sigma - \beta_1} \leq \frac{6/7}{\sigma - 1} + L \quad \text{for} \quad 1 < \sigma \leq 1 + \frac{1}{21C^L}
\]  

(8)

For in case (i) we have \(\delta_1 = \delta_2 = 0\) and \(3/4 \leq 6/7\); and in case (ii) we have \(\delta_1, \delta_2 \leq 1\) and \(|y_1| \geq 3(\sigma - 1)\) hence
\[
\frac{4(\sigma - 1)}{(\sigma - 1)^2 + \gamma_1^2} + \frac{\sigma - 1}{(\sigma - 1)^2 + 4\gamma_1^2} \leq \left(\frac{4}{10} + \frac{1}{37}\right)\frac{1}{\sigma - 1} < \frac{3}{7}\frac{1}{\sigma - 1}
\]

as claimed (note that to apply 7) we must verify \(1 + \frac{1}{21C^L} \leq 2\) for which we may need to raise the \(\mathcal{L} \gg n\) constant). Thus, putting \(\sigma = 1 + 1/21C^L\) in (8) we deduce in case (i), (ii) that
\[
\beta_1 \leq 1 - \frac{2}{399C^L} < 1 - \frac{1}{200C^L}
\]

so putting \(c_1 = 1/400C\) it follows that we are in case (iii).
In case (iii) we must proceed differently: putting aside our fixed zero \((\rho_1, \chi_1)\) for the moment we study zeros of Dedekind zeta functions \(\zeta_E\) of extension fields \(E\) of \(K\) which satisfy

\[
[E:K] \leq 4 \quad \text{and} \quad d_E \leq Q^4
\]

and indeed will prove the **CLAIM**: If \(E\) satisfies the condition (9) then in the region

\[
R: \quad |t| < \frac{1}{7C_L}, \quad \sigma \geq 1 - \frac{1}{200C_L}
\]

there is at most one zero of \(\zeta_E(s)\); this zero is real and simple (if it exists).

Again we start from the basic inequality (4), applied this time to \(\zeta_E\), with \(s = \sigma\) real and between 1 and 2. In place of (6) we have the simple inequality \(\frac{-r'_E}{\zeta_E}(\sigma) \geq 0\) and applying the conditions (9) we get

\[
\sum_{\rho \in \mathcal{L}_E} \Re \frac{1}{\sigma - \rho} \leq \frac{1}{\sigma - 1} + 2 \log Q + O(n)
\]
since \(|s|\) is bounded and \([E:Q] \ll n\). By (17) of 52 this simplifies to

\[
\sum' \Re \frac{1}{\sigma - \rho} \leq \frac{1}{\sigma - 1} + 2C_L, \quad 1 \leq \sigma \leq 2
\]

where \(\sum'\) denotes summation over any subset of \(\mathcal{L}_E\), and where we use \(\log Q \leq \mathcal{L}\) and \(\mathcal{L} \gg n\) to eliminate the \(O(n)\) term (possibly raising the \(\gg\) constant again).
To deduce the claim from (10) we again consider 3 cases:

a) \( \tau_E \) has a complex zero in \( R \)

b) \( \tau_E \) has a real double zero in \( R \)

c) \( \tau_E \) has two distinct real zeros in \( R \)

and show that each leads to a contradiction, by showing first that in each case there is a zero \( \rho = \beta + i\gamma \) of \( \tau_E \) satisfying

\[
\frac{8}{5} \frac{1}{\sigma - \beta} < \frac{1}{\sigma - 1} + 2CL \quad \text{for} \quad 1 + \frac{2}{7CL} \leq \sigma \leq 2 \tag{11}
\]

For in case (a) if \( \rho = \beta + i\gamma \) is the zero assumed in \( R \) then
\( \rho = \beta - i\gamma \) is also a zero of \( \tau_E \), by (10) of §2, hence restricting the sum (10) to these two (different) zeros gives

\[
2 \frac{\sigma - \beta}{(\sigma - \beta)^2 + \gamma^2} \leq \frac{1}{\sigma - 1} + 2CL .
\]

However from \( |\gamma| < \frac{1}{7CL} \leq \frac{1}{2}(\sigma - 1) \leq \frac{1}{2}(\sigma - \beta) \) we have

\[
2 \frac{\sigma - \beta}{(\sigma - \beta)^2 + \gamma^2} > 2 \frac{1}{\sigma - 1} + \frac{1}{4} \frac{1}{\sigma - \beta} = \frac{8}{5} \frac{1}{\sigma - \beta} ,
\]

hence (11). And in case (b) if \( \rho = \beta + i\gamma \) is the double zero assumed in \( R \) then (10) gives directly

\[
\frac{2}{\sigma - \beta} \leq \frac{1}{\sigma - 1} + 2CL \quad \text{which is more than enough.}
\]

Finally in case (c) let \( \beta < \beta' \) be the two real zeros assumed in \( R \): by (10) we have

\[
\frac{1}{\sigma - \beta} + \frac{1}{\sigma - \beta'} \leq \frac{1}{\sigma - 1} + 2CL ,
\]

and by \( \beta < \beta' \) we have

\[
\frac{2}{\sigma - \beta} < \frac{1}{\sigma - \beta} + \frac{1}{\sigma - \beta'} ,
\]

so (11) again follows.

Putting \( \sigma = 1 + \frac{2}{7CL} \) in (11) and noting that this is \( \leq 2 \) by \( L \gg n \) it follows that

\[
\beta \leq 1 - \frac{2}{385CL} < 1 - \frac{1}{200CL} < \frac{1}{2}
\]

which contradicts the hypothesis that \( \beta + i\gamma \) was in \( R \). Thus the claim is proved.
We can now return to our fixed zero \((\rho_1, \chi_1)\) in case (iii) i.e. \(\chi_1\) is real valued and \(\rho_1\) is in the region \(R\) of the claim; and, indeed since \((\rho_1, \chi_1)\) was arbitrary, we already know that any other zero \((\rho, \chi)\) in the region of the statement of the lemma must be in case (iii).

Suppose first that \(\chi_1 \neq 1\): then \(\ker \chi_1\) has index 2 in \(I(\chi_1)\) hence the class field \(E/K\) to \(\ker \chi_1\) has \([E:K] = 2\). Moreover \(d_E \leq Q^4\) also holds (for \(d_{E/K} = \chi_1\)) by the conductor-discriminant formula hence by the tower formula for discriminants we have
\[
d_E = d_{K(\chi_1)} = d_K \chi_1 \leq Q^2\)
so since (by p. 230 of Lang [18])
\[
\zeta_E(s) = \zeta_K(s) L(s, \chi_1)
\]
the claim shows that \(\rho_1\) is the only zero of \(L(s, \chi_1)\) in \(R\), that \(\rho_1\) is a real simple zero of \(L(s, \chi_1)\), and that \(\zeta_K(s)\) has no zeros in \(R\).

Turning now to \(\chi_1 = 1\) we can apply the claim directly to \(E = K\) showing again that \(\rho_1\) is a real simple zero and the only zero of \(L(s, \chi_1)\) in \(R\). Moreover the construction of the previous paragraph (applied to an arbitrary real-valued character \(\psi\) with \(d_\psi \leq Q\) in place of the \(\chi_1\) there) shows that no \(L(s, \psi)\) with \(\psi \neq 1\), \(d_\psi \leq Q\) can have a zero in \(R\) so that the lemma is proved in case \(\chi_1 = 1\).

Returning to \(\chi_1 \neq 1\) it now remains only to rule out the possibility of a zero \((\rho_2, \chi_2)\) for a real-valued character \(\chi_2 \neq 1, \chi_1\), where \(d_{\chi_2} \leq Q\) holds. So suppose this happens: then \(\ker \chi_1 \cap \ker \chi_2\) has index 4 in \(I(\chi_1, \chi_2)\) hence the corresponding class field
E/K has \([E:K] = 4\). Letting \(\chi_3\) be the primitive character which induces \(\chi_1\chi_2\) we have

\[
\zeta_E(s) = \zeta_K(s)L(s,\chi_1)L(s,\chi_2)L(s,\chi_3)
\]

so since \(d_E \leq Q^4\) (for by the conductor-discriminant formula
\(d_{E/K} = \delta_{\chi_1} \delta_{\chi_2} \delta_{\chi_3}\), and since \(\chi_3\) induces \(\chi_1\chi_2\) we have \(\delta_{\chi_3}\) dividing \(\delta_{\chi_1} \delta_{\chi_2}\). Thus by the tower formula for discriminants

\[
d_E = d_{E/K}^N(\delta_{\chi_1} \delta_{\chi_2} \delta_{\chi_3}) \leq d_{K}^N(\delta_{\chi_1} \delta_{\chi_2} \delta_{\chi_3})^N(\delta_{\chi_1} \delta_{\chi_2} \delta_{\chi_3}) = d_{\chi_1}^2 d_{\chi_2}^2 \leq Q^4
\]

applying

the claim shows that \(\rho_2\) cannot exist, and completes the proof of lemma 2.

It may be useful to observe that the use of class field theory above can be avoided: for the functions we have denoted \(\zeta_E(s)\) can be constructed simply by taking appropriate products of L-series (i.e. without knowing they are zeta functions of fields) and the relevant properties verified directly (if tediously). The approach chosen has, however, the advantage of motivating the constructions (and of brevity).

We also need a \(K\)-uniform version of Linnik's Density Lemma, which we state as

**Lemma 3.** Let \(Q \geq 1, T \geq 1\) satisfy \(Q^T > 1\) and put \(\mathcal{L} = \log QT^N\). If \(|\tau| \leq T, \tau \in \mathbb{R}\) and \(d_\chi \leq Q\) (with \(\chi\) primitive) then the number (counting multiplicities) of zeros of \(L(s,\chi)\) in the disc

\[
|s - (1+i\tau)| \leq r
\]

is

\[
\ll r^{\mathcal{L}}
\]

uniformly for \(\frac{1}{2} \mathcal{L}^{-1} \leq r \leq 1\), where \(c_2 > 0\).
Proof: We employ the same method as in lemma 2 and begin by applying (4) with \((s,\psi) = (1+r+it, \chi)\) where \(r, t, \chi\) satisfy the above hypothesis obtaining

\[
-\operatorname{Re} \frac{L'}{L}(1+r+it, \chi) + \sum_{\rho \in \mathcal{L}(\chi)} \operatorname{Re} \frac{1}{1 + r + it - \rho}
\]

\[
\leq \delta \frac{r}{r^2 + t^2} + \frac{1}{2} \log d_{\chi} + \frac{n}{2} \log |1 + r + it| + O(n)
\]

\[
\leq \frac{1}{r} + \frac{1}{2} \log QT^n + O(n) \ll \mathcal{L} \text{ since } \frac{1}{r} \ll \mathcal{L} \text{ and } n \ll \mathcal{L}.
\]

And we also apply (4) with \((s,\psi) = (1+r,1)\) so

\[
-\zeta' K(1+r) + \sum_{\rho \in \mathcal{L}_K} \operatorname{Re} \frac{1}{1 + r - \rho} \leq \frac{1}{r} + \frac{1}{2} \log d_{\chi} + \frac{n}{2} \log |r+1| + O(n)
\]

\[
\leq \frac{1}{r} + \frac{1}{2} \log Q + O(n) \ll \mathcal{L} \text{ as before.}
\]

We propose to add these two inequalities but first observe that by (16) of §2 we have

\[
-\zeta' K(1+r) - \operatorname{Re} \frac{L'}{L}(1+r+it, \chi) = \sum \frac{\Lambda(a)}{N(a)^{1+r}} (1+\operatorname{Re} \chi(a)N(a)^{-it})
\]

\[
\geq \sum \frac{\Lambda(a)}{N(a)^{1+r}} (1 - |\chi(a)N(a)^{-it}|) \geq 0 \text{ since } |\chi(a)N(a)^{-it}| = 1 \text{ or } 0
\]

and that, by (17) of §2, we may omit any of the terms in the sums over \(\rho\). Thus we get

\[
\sum_{\rho \in \mathcal{L}(\chi), |\rho-(1+it)| \leq r} \operatorname{Re} \frac{1}{1 + r + it - \rho} \ll \mathcal{L}
\]

where, letting \(\rho\) be one of the zeros in the sum, (i.e. \(|\rho-(1+it)| \leq r\)) we have

\[
|1 + r + it - \rho| \leq r + |1 + it - \rho| \leq 2r
\]

from which
\[
\text{Re} \frac{1}{1 + r + i \tau - \rho} = \frac{1 + r - \text{Re} \rho}{|1 + r + i \tau - \rho|^2} \geq \frac{r}{(2r)^2} = \frac{1}{4r}
\]

which combines with (12) to give
\[
\frac{1}{4r} \sum_{\rho \in \mathcal{L}(\chi)} 1 \ll \mathcal{L} \quad \text{and} \quad \left| \rho - (1 + i \tau) \right| \leq r
\]
proves the lemma. Indeed it is clear that any fixed \( c_2 > 0 \) will do.

Together lemmas 2 and 3 give some limitations on how the zeros of L-series can be distributed in the critical strip \( 0 \leq \sigma \leq 1 \). Our next lemma gives some different kind of information: roughly, it says that if \( L(s, \chi) \) has zeros "too close" to some complex number \( s \) with \( \sigma > 1 \) then, in light of (17) of §2, \( \sum \frac{\Lambda(a) \chi(a)}{a N(a)^s} \) must be "rather large". More precisely we have

**Lemma 4** Let \( Q \geq 1, T \geq 1 \) satisfy \( Q^T \gg 1 \) and put \( \mathcal{L} = \log QT^n \).

If \( |\tau| \leq T, \tau \in \mathbb{R} \) and \( d \chi \leq Q \) then

\[
\left| \frac{L'(s, \chi)}{L(s, \chi)} + \frac{\delta \chi}{s-1} \sum_{\rho \in \mathcal{L}(\chi)} \frac{1}{s-\rho} \right| \ll \mathcal{L}
\]

holds, uniformly, for \( |s - (1 + i \tau)| \leq \frac{1}{2} \).

**Proof:** Letting \( \tau, \chi, s \) satisfy the above hypothesis we again use the same ideas, however this time we cannot eliminate \( b_\chi \) in (13) of §2 simply by taking real parts, since our conclusion is not just about the real part. Instead we will use the same identity (13) of §2 at a value where we can estimate the \( L'/L \) term directly, and subtract this to eliminate the \( b_\chi \) term.

We begin then with (13) of §2 for \( 2 + i \tau \) i.e.
Proceeding as at the beginning of the proof of lemma 2 we find that the right side of this identity is

\[ \ll L \]

since Stirling's formula applies not just to real parts, and since \( n \ll L \). On the left side of our identity we observe that, by (17) of §2, we have

\[
\left| \frac{-L'}{L}(2+i\tau, \chi) \right| \leq \sum \frac{\Lambda(a)}{\alpha N(a)^2} |\chi(a)N(a)^{-1}\tau| \leq \sum \frac{\Lambda(a)}{\alpha N(a)^2}.
\]

and since for each rational prime power \( p^m \) there are \( \leq \) integral ideals \( a \) with \( N(a) = p^m \), we have

\[
\sum \frac{\Lambda(a)}{\alpha N(a)^2} \leq n \sum \frac{\Phi(m)}{m^2} \ll n
\]

hence that

\[
\left| \frac{-L'}{L}(2+i\tau, \chi) \right| \ll n \quad (13)
\]

so our identity yields, on using \( n \ll L \) again:

\[
\left| -\frac{b}{\chi} + \sum_{\rho \in \rho_{\chi}} \left( \frac{1}{2+i\tau-\rho} + \frac{1}{\rho} \right) \right| \ll L \quad (14)
\]
Before applying this we observe that by (12) of §2 we have

\[ \sum_{\rho} \Re \frac{1}{2 + i\tau - \rho} \leq -b \chi + \sum_{\rho} \left( \frac{1}{2 + i\tau - \rho} + \frac{1}{\rho} \right) \ll \mathcal{L} \]

where also

\[ \Re \frac{1}{2 + i\tau - \rho} = \frac{2 - \Re \rho}{|2 + i\tau - \rho|^2} \geq \frac{1}{|2 + i\tau - \rho|^2} \]

hence

\[ \sum_{\rho} \frac{1}{|2 + i\tau - \rho|^2} \ll \mathcal{L} \quad (15) \]

Of course (15) could have been obtained from (4) and (13); however, we also need (14).

Going back to (13) of §2 again, this time with \( s \) as in the statement of the lemma, we find

\[ \frac{L'}{L}(s, \chi) + \frac{\delta}{s-1} \sum_{\rho} \left( \frac{1}{s-\rho} + \frac{1}{\rho} \right) + b \chi = -\log A \chi - \frac{\delta}{s} - \frac{\Gamma'}{\Gamma}(s) \]

Now since \( |s - (1+i\tau)| \leq \frac{1}{2} \) and \( |\tau| \leq T \), where \( T \geq 1 \), applying Stirling's formula again results in

\[ \left| \frac{\Gamma'}{\Gamma}(s) \right| \leq \frac{n}{2} \log T + O(n) \]

and since also \( \frac{1}{s} \ll 1 \), the right hand side of our identity is \( \ll \mathcal{L} \).

Using this and adding (14), as suggested at the beginning, leaves

\[ \left| \frac{L'}{L}(s, \chi) + \frac{\delta}{s-1} - \sum_{\rho} \left( \frac{1}{s-\rho} - \frac{1}{2 + i\tau - \rho} \right) \right| \ll \mathcal{L} \]
which can be rewritten as

\[
\left| \frac{L'(s,L)}{L(s)} + \frac{\delta}{s-1} \right| \leq \sum_{\rho-(1+i\tau)} \frac{1}{s-\rho} \left| \rho-(1+i\tau) \right| < 1 \leq \sum_{\rho-(1+i\tau)} \frac{1}{2+2i\tau-\rho}.
\]

\[
(16)
\]

so we are left with estimating the two sums over \( \rho \) which we denote \( S_1, S_2 \) temporarily: it is for this purpose that we need (15).

Let \( \rho \) be as in \( S_1 \), namely \( |\rho-1-i\tau| < 1 \): then

\[
2+2i\tau-\rho \leq 1 + |1+i\tau-\rho| < 2 \text{ hence } \frac{1}{2+2i\tau-\rho} < \frac{2}{|2+2i\tau-\rho|^2} \text{ so we have }
\]

\[
\sum_1 \sum_{\rho-(1+i\tau)} \frac{1}{|2+2i\tau-\rho|^2} \ll \mathcal{L}, \text{ by (15)}.
\]

Finally let \( \rho \) be as in \( S_2 \), namely \( |\rho-1-i\tau| \geq 1 \): then we have

\[
|2+i\tau-s| \leq 1 + |1+i\tau-s| \leq 3/2
\]

and

\[
\frac{1}{2} \leq |1+i\tau-\rho| - \frac{1}{2} \leq |1+i\tau-s| + |s-\rho| - \frac{1}{2} \leq |s-\rho|
\]

hence

\[
2+2i\tau-\rho \leq |2+i\tau-s| + |s-\rho| \leq \frac{3}{2} + |s-\rho| \leq 4 |s-\rho|
\]

From the first and last of these we find

\[
\sum_2 = \sum_{|\rho-1-i\tau| \geq 1} \frac{|2+i\tau-s|}{|s-\rho||2+i\tau-\rho|} \leq \sum_{|\rho-1-i\tau| \geq 1} \frac{4 \cdot 3/2}{|2+2i\tau-\rho|^2} \ll \sum_2 \frac{1}{\rho|2+2i\tau-\rho|^2}
\]

\[ \ll \mathcal{L} \], by (15) again. Together with (16) the lemma is proved.

**COROLLARY.** Let \( Q \geq 1, T \geq 1 \) satisfy \( Q^T > 1 \) and put \( \mathcal{L} = \log QT^n \).

If \( |\tau| \leq T \) and \( d\chi \leq Q \) then the number of zeros \( \rho \) of \( L(s,\chi) \) satisfying \( |\text{Im} \rho - \tau| \leq 1 \) is \( \ll \mathcal{L} \), while for the zeros outside the rectangle we have \( \sum \frac{1}{|\text{Im} \rho - \tau| \geq 1} \ll \mathcal{L} \).

**PROOF:** This is a consequence of (15), since we have

\[
\frac{1}{|2+i\tau-\rho|^2} = \frac{1}{(2-\text{Re} \rho)^2 + (\text{Im} \rho - \tau)^2} \geq \frac{1}{4 + (\text{Im} \rho - \tau)^2}
\]

\[
\geq \begin{cases} 
\frac{1}{3} & \text{if } |\text{Im} \rho - \tau| \leq 1 \\
\frac{1}{5(\text{Im} \rho - \tau)^2} & \text{if } |\text{Im} \rho - \tau| > 1
\end{cases}
\]

It can also be deduced from lemma 3, however the purpose (and strength) of lemma 3 is that it is sharper for small values of \( r \).

Finally, before concluding this §, we mention (mostly without proof, although most of the ingredients are here) a few further results which suggest the appropriateness of the above considerations: first, if in the first result of the corollary we let \( \tau \) take on all the integer values with \( |\tau| \leq T \) then we find that the number of zeros of \( L(s,\chi) \) which satisfy \( |\text{Im} \rho| \leq T \) is \( \ll T\mathcal{L} = T \log QT^n \), and that this holds for every \( \chi \) with \( d\chi \leq Q \).
In fact, using these estimates, as in chapter 16 of Davenport [4], the same number of zeros can be shown to be equal to

\[ \frac{1}{\pi} \log d \chi T^n - n \left( \frac{1 + \log 2\pi}{\pi} \right) T + O(\log d \chi T^n) \]

which provides some assurance that the formation \( L = \log QT^n \) is appropriate; in particular \( T^n \) is needed, not just \( T \).

Second, in the notation of lemma 2, we have for \( d \chi \leq Q \)

\[ b_\chi = \begin{cases} \frac{1}{1-\rho_1} + O(L^2) & \text{if } \chi = \chi_1 \text{, whenever } (\rho_1, \chi_1) \text{ exists} \\ 0(L^2) & \text{otherwise} \end{cases} \]

which is an easy consequence of (14), with \( \tau = 0 \), and the corollary. Since, as far as is known, \( 1/(1-\rho_1) \) can be considerably larger than \( L^2 \) this shows that the techniques for eliminating \( b_\chi \) are indeed necessary.

Lastly, on a more negative note, we observe that in spite of the fact that all the results of this § depend on the identities (13) and (14) of §2, the full content of these identities has hardly been utilized, for in each application we used only rather weak consequences, obtained by throwing away those parts we could not handle. Indeed, Odlyzko [21] has been able to employ (14) of §2, with \( \chi = 1 \), to improve on the Minkowski lower bounds for discriminants, an indication that more information may be obtainable.

§4 The exceptional zero

Given \( Q, T \geq 1 \) as in lemma 2, we call the zero \( (\rho_1, \chi_1) \) of
lemma 2 the exceptional zero for \( Q, T \) (and \( K \)), when it exists. Conjecturally the exceptional zero does not exist, of course; however, in the present state of knowledge, we must allow this possibility and treat the cases when it exists and when it doesn't separately. In this § we want to adapt the results of Stark [27], which show that when \((\rho_1, \chi_1)\) exists then \(\rho_1\) is not too close to 1 (recall that \(\rho_1\) is real < 1).

First we want to emphasize, however, that the possibility \(\chi_1 = 1\) (the trivial character) has not been ruled out: indeed if we could eliminate this possibility then we would be able to prove that there are no exceptional zeros (by class field theoretic arguments as at the end of the proof of lemma 2). This is a deviation from the situation when \(K = \mathbb{Q}\) for then it is easy to see that \(\chi_1 \neq 1\) (for

\[
(1-2^{-s})\zeta_\mathbb{Q}(s) = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m^s} \text{ for } s > 0
\]

as on p. 157 of Lang [18]; since the series is alternating for real values of \(s\), \(0 < s < 1\), it follows that \(\zeta_\mathbb{Q}(s) < 0\) for these values of \(s\).

Turning now to the results of Stark [27] we begin by noting that his results are stated in a slightly different form, which is better adapted to this discussion, and also that they are entirely effective (indeed, he gives the constants numerically), so we content ourselves with a sketch of his methods and adapt them to our purposes later.

We consider only zeta functions of fields \(E \neq \mathbb{Q}\) and make the following definition: we say that \(\beta_p\) is the bad zero of \(E\) if
\[ \beta_E \text{ is real, } 1 - \frac{1}{4 \log d_E} < \beta_E < 1 \text{ and } \tau_E(\beta_E) = 0; \quad (1) \]

just as in our lemma 2 Stark proves that \( \beta_E \) is a simple zero of \( \tau_E(s) \) and the only zero of \( \beta_E(s) \) satisfying (1), provided that it exists at all. Before continuing we note that the term "bad" is used only to avoid confusion with our previous definition and that the restriction \( E \neq \mathbb{Q} \) is entirely natural in view of the second paragraph of this § (indeed it is made only to avoid bothering with \( \log d_\mathbb{Q} = 0 \), for \( E \neq \mathbb{Q} \) implies \( d_E > 1 \), as we know).

The first result we need is the relation between \( \beta_E \) and the residue \( \kappa(E) \) of \( \tau_E(s) \) at \( s = 1 \), namely

\[
\text{if } E \neq \mathbb{Q} \text{ then } \kappa(E) = \begin{cases} 
(1-\beta_E), & \text{if } \beta_E \text{ exists} \\
\frac{1}{\log d_E}, & \text{if not}
\end{cases}
\quad (2)
\]

which is Stark's lemma 4, or follows also from p. 323 of Lang [18] (this is an important step in the proof of the Brauer-Siegel theorem).

However the main result is a lower bound for \( (1-\beta_E) \) and is deduced from the following result which is Stark's elaboration of a theorem of Heilbronn [13]:

**Theorem 1.** Let \( L/F \) be a finite galois extension of number fields and let \( \lambda \) be a simple zero of \( \zeta_L(s) \). Then there is a (unique) intermediate field \( E \) of \( L/F \) so that for any intermediate field \( E' \) of \( L/F \) we have \( \zeta_{E'}(\lambda) = 0 \) if and only if \( E' \supseteq E \). Moreover \( E/F \) is a cyclic extension, and if \( \lambda \) is real then \([E:F] = 1 \) or 2.
This remarkable result replaces Artin's conjecture on non-abelian L-series (but only for simple zeros) and is proved by character theory.

Application of theorem 1 to the problem of bad zeros depends on the following simple result on the behaviour of discriminants: if \( E_1, \ldots, E_m \) are number fields with composite \( E \) then

\[
\frac{1}{d_E} \leq \prod_{i=1}^{m} \frac{1}{d_{E_i}}
\]

which Stark proves as lemma 6. Putting (3) and theorem 1 together appropriately allows an inductive argument from which follows the main result:

**Theorem 1'** If the number field \( E \) has a bad zero \( \beta_E \) then

\[
(1-\beta_E) \gg \min\left(\frac{1}{c_E \log d_E}, \frac{1}{d_E^{1/n_E}}\right)
\]

where \( n_E = [E: \mathbb{Q}] \) and where \( c_E = 1 \) if there is a tower \( Q = E_0 \subseteq E_1 \subseteq \ldots \subseteq E_t = E \) of fields with every \( E_i/E_{i-1} \) normal, while \( c_E = n_E! \) in general.

This is the 'em l' of Stark [27], or rather this is what he proves: he then states his theorem as a lower bound for \( \kappa(E) \) ("explicit form of the Brauer-Siegel theorem") by combining the above statement with (2). For our applications the main defect of theorem 1' is the appearance of \( n_E! \) in the general case (interestingly this is also the case for Stark's applications, as he points out), a circumstance which will be discussed further when it arises. For now we observe that \( n_E! \) occurs in theorem 1' from the need to take
normal closures (so theorem 1 applies): we can estimate the discriminant of the normal closure by (3) (and $n_E!$ appears) and must use this estimate to ensure that the bad zero of $E$ remains bad for the normal closure, hence is simple (so theorem 1 applies). It is therefore not too surprising that the appearance of $n_E!$ can be eliminated on the Artin conjecture (which theorem 1 replaced) as Stark proves in his theorem 4; namely we have

**THEOREM 1"**. If Artin $L$-series are entire (except for a pole at $s = 1$ for the trivial character) then theorem 1' holds with $c_E = n_E$ in the general case.

Finally we want to relate this to our previous situation so, suppose $K$ is again fixed, that we are given $Q, T$ as in lemma 2, and that the exceptional zero $(p_1, x_1)$ exists: let $E$ be the class field to ker $\chi_1$. Then we have $[E:K] = 1$ or 2 (since $\chi_1$ is real-valued) and $d_E = d_K d_{x_1} \leq Q^2$ (or even $Q$ if $\chi_1 = 1$). As in the proof of lemma 2 we have

$$\zeta_E(s) = \begin{cases} 
\zeta_K(s)L(s, x_1) & x_1 \neq 1 \\
\zeta_K(s) & x_1 = 1
\end{cases}$$

hence $p_1$ is a simple real zero of $\zeta_E(s)$ with

$$1 > p_1 > 1 - 2c_1 / \log QT^n > 1 - c_1 / \log d_E$$

by lemma 2 (and $d_E \leq Q^2$). Since our constant $c_1$ is $< \frac{1}{4}$ (by the proof of lemma 2, which proof is also the source of the 1/4) it follows
from (2) that $\rho_1$ is the bad zero $\beta_E$ of $E$, which exists because $\rho_1$ does. Thus by theorem 1' we get a lower bound for $1 - \rho_1$ which we state in the notation of §3 as

$$1$$

**COROLLARY.** Suppose $Q \geq 1, T \geq 1$ satisfy $Q^T \gg 1$ and put $\mathcal{L} = QT^n$. If the exceptional zero $(\rho_1, \chi_1)$ for $Q, T$ exists then

$$(1 - \rho_1) \gg \min\left(\frac{1}{c_{K,1} \log Q}, \frac{1}{Q^{1/n}}\right) \approx \min\left(c_{K,1}^{-1} \mathcal{L}^{-1}, e^n\right)$$

where $c_{K,1} = 1$ if there is a tower $\mathcal{Q} = K_0 \subset K_1 \subset \ldots \subset K_t = K$ of fields with every $K_i/K_{i-1}$ normal, and $c_{K,1} = (2n)!$ in general.

**Proof:** It remains only to observe that from $d_E \leq \mathcal{Q}^{[E:K]}$ follows $d_E^{1/n_E} \leq Q^{1/n}$, that if there is a "normal tower" to $K$ then there is also one to $E$ (since $E/K$ is itself normal), and (for the transition from $Q$ to $\mathcal{L}$) that $T \geq 1$. 
CHAPTER II

A SPECIAL ARITHMETIC SUM

The sum in question is

\[ \sum_{a} 1/N(a) \]

where \( a \) runs through integral ideals of a field and is restricted by \( N(a) \) being in a certain range. Indeed \( N(a) \) will be restricted in two ways and we are interested in a lower bound in one case and in an upper bound in the other. The point, in both cases, is that these bounds can be obtained by replacing the given sum by a suitable average which can be estimated more effectively. Since we do not need an asymptotic equality for the original sum but rather a lower (or upper) bound of the correct order of magnitude we can, in this way, obtain estimates which are valid for somewhat more restricted ranges than is usual. Finally we apply these results to estimate a similar sum over prime ideals.

§1 A lower bound.

It is notationally appropriate to consider sums of integral ideals of a field \( E \), because the results will ultimately be applied to quadratic extensions of our original field \( K \), as well as to \( K \) itself (exactly as in the last paragraph of chapter I); we use also the "bad zero" of \( I \& 4 \) because it turns out that better estimates can be obtained when it exists, a phenomenon which has been exploited in the present context by Bombieri [2] (for lemma C of his chapter 6) in the special case \( K = \mathbb{Q} \). Indeed Bombieri even uses the same
average as used here, but only to ensure absolute convergence of an 
integral.

The average in question is

$$\sum_{Na<x} \frac{1}{Na(1-Na)}^k$$

where $k$ is a positive integer which we choose depending on the field $E$. The estimation of such sums is based on the familiar integral formula

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{y^s}{s(s+1)...(s+k)} \, ds = \begin{cases} 0 & 0 < y \leq 1 \\ \frac{1}{k!(1-\frac{1}{y})} & y > 1 \end{cases}$$

valid for real $c > 0$ and integral $k \geq 1$. We include a proof of this for convenience: most of the work is done by the simple inequality

$$|s(s+1)...(s+k)| \geq |s|^{-k} \geq \left(\frac{1}{2}\right)^{k+1} \text{ for } |s| \geq U \geq 2k.$$ (2)

First, if $y \leq 1$, consider the positively oriented rectangle with vertices at $c-iU$, $U-iU$, $U+iU$, $c+iU$ where $U > 0$ is large (we will let $U \to \infty$): using $y \leq 1$ and 2) shows that the integral on the horizontal edges is in absolute value

$$\leq \frac{1}{2\pi} \int_{U-\frac{1}{2}U}^{U} \frac{y^c}{1-\frac{1}{U}k+1} \, ds \ll \frac{y^c}{U^{k+1}} \int_{c}^{U} \, ds \ll U^{-k}$$

and that the integral on the vertical edge $\sigma = U$ is in absolute value
so, since there are no poles of the integrand inside our rectangle, we get our contention on letting $U \to \infty$.

Similarly, if $y > 1$, we consider the rectangle with vertices at $-U - iU$, $c - iU$, $c + iU$, $-U + iU$ where $U > 0$ again will tend to $\infty$. Estimating as above, we find that the integrals along the horizontal edges and along the vertical edge $\sigma = -U$ tend to 0 as $U \to \infty$, hence our integral is the sum of the residues of the integrand in the half plane $\sigma < c$: there is a simple pole at $s = -j$ for each integer $j$ with $0 \leq j \leq k$ and the residue there is $\frac{(-1)^j y^{-j}}{j!(k-j)!}$ so the sum of the residues is

$$\sum_{j=0}^{k} \frac{(-1)^j y^{-j}}{j!(k-j)!} = \frac{1}{k!} \left(1 - \frac{1}{y}\right)^k$$

as claimed.

**Lemma 5.** We have, for $x \geq 1$

$$\left| \sum_{Na < x} \frac{1}{Na} (1-Na) \frac{n_E}{x} - \kappa(E)(\log x - \sum_{j=1}^{k} \frac{1}{j}) - \kappa_0(E) \right| \leq e^{-\theta(n_E)} e^{n_E \frac{3}{8} x 1^2}$$

Proof: Let real $a \geq 1$, integer $k \geq 1$ be specified later: we show that a rather stronger estimate is true for suitably restricted $k$, however the stated result will be adequate later, and is chosen for its simplicity. Applying (1) with $c = 1/a$ and $y = x/Na$ gives

$$\sum_{Na < x} \frac{1}{Na} (1-Na) \frac{k}{x} = \sum_{a} \frac{1}{Na} 2\pi i \int_{\frac{1}{a-i\infty}}^{\frac{1}{a+i\infty}} \frac{s}{Na} \frac{ds}{s(s+1)\ldots(s+k)}$$
where the integral converges absolutely by \( k \geq 1 \). Since the Dirichlet series for \( \zeta_E(s+1) \) also converges absolutely for \( \sigma = \frac{1}{a} > 0 \) this becomes

\[
\sum_{Na \leq x} \frac{1}{N \alpha(1 - \frac{Na}{x})^k} = \frac{1}{2\pi i} \int_{\frac{1}{a} - i \infty}^{\frac{1}{a} + i \infty} \zeta_E(s+1) \frac{x^s}{s(s+1) \cdots (s+k)} \, ds, \quad x > 0 \quad (3)
\]

and we propose next to move the line of integration from \( \sigma = \frac{1}{a} \) to \( \sigma = -\frac{1}{2} \) (it is actually better to move it to \( \sigma = -1 - \frac{1}{a} \), but this requires a little more effort) for which purpose we consider the positively oriented rectangle with vertices at \( -\frac{1}{2} - iU, \frac{1}{a} - iU, \frac{1}{a} + iU, -\frac{1}{2} + iU \) where \( U > 0 \) will again tend to \( \infty \). We estimate the integrals along the horizontal edges qualitatively to show they tend to 0: by lemma 1 we have

\[
|\zeta_E(s+1)| \leq e^{-\frac{U}{2}} a d_E \left( \frac{1}{2} \right)^{\frac{1}{a}} (1+U)^{-\frac{1}{2} - \frac{1}{a}} \ll (U)^{\frac{1}{2} - \frac{1}{a}}
\]

on these edges, hence using (2) again the integrals along the horizontal edges are in absolute value

\[
\ll \int_{-\frac{1}{2}}^{\frac{1}{a}} \frac{1}{y^{2} n_E \left( \frac{1}{a} + \frac{1}{2} \right)} x^{\sigma} k! d\sigma \ll \frac{1}{y^{2} n_E \left( \frac{1}{a} + \frac{1}{2} \right)} -k-1
\]

which tends to 0, provided that \( k + 1 > \frac{1}{2} n_E \left( \frac{1}{a} + \frac{1}{2} \right) \). Moreover, inside our rectangle there is only a double pole at \( s = 0 \), with residue

\[
\kappa(E) \log x - \kappa(E) \sum_{j=1}^{k} \frac{1}{j} + \kappa_0(E)
\]
as follows from the Laurent expansion, (11) of I§1, for \( \zeta_E(s+1) \).

Thus letting \( U \to \infty \), (3) becomes

\[
\sum \frac{1}{Na} \left( 1 - \frac{1}{x} \right)^k - \kappa(E) \log x + \kappa(E) \sum_{j=1}^{k} \frac{1}{j} - \kappa_0(E)
\]

\[= - \frac{1}{2\pi i} \int_{-\frac{1}{2} + i\infty}^{-\frac{1}{2} - i\infty} \zeta_E(s+1) \frac{x^s}{s} \frac{k! ds}{(s+1)...(s+k)}
\]

provided that \( k + 1 > \frac{1}{2} n_E \left( \frac{1}{a} + \frac{1}{2} \right) \).

It thus remains only to estimate the integral in (4); from lemma 1 again

\[
|\zeta_E(s+1)| \leq e^{-\rho E} a E d_E^2 \left( 1 + \frac{1}{a} \right) \left( 1 + |t| \right)^{1/2} \frac{1}{2 n_E \left( \frac{1}{2} + \frac{1}{a} \right)}
\]

on \( \sigma = -\frac{1}{2} \)

so the integral is in absolute value

\[
\leq k! e^{-\rho E} a E d_E^2 \left( 1 + \frac{1}{a} \right) \frac{1}{2 n_E \left( \frac{1}{2} + \frac{1}{a} \right)} \int_{-\infty}^{\infty} \frac{(1+|t|)^{1/2}}{(j-\frac{1}{2})^2 + t^2} dt
\]

where the final integral is

\[
\leq 2 \int_{0}^{\infty} \frac{1}{2 n_E \left( \frac{1}{2} + \frac{1}{a} \right)} \frac{1}{k+1} \frac{1}{(1+t^2)^{1/2}} dt \ll \int_{0}^{1} 0(k) \, dt + \int_{1}^{\infty} 0(n_E) \frac{1}{2 n_E \left( \frac{1}{2} + \frac{1}{a} \right)}^{-k+1} \, dt
\]

\[
\ll e^{\rho E} + \frac{e^{\rho E}}{k - \frac{1}{2} n_E \left( \frac{1}{2} + \frac{1}{a} \right)} \text{ provided } k > \frac{1}{2} n_E \left( \frac{1}{2} + \frac{1}{a} \right).
\]
It is now clear which values of $k$ give a reasonable estimate, so putting $k = n_{E}$ (which is $> \frac{1}{2} n_{E} (\frac{1}{2} + \frac{1}{a})$ by $a \geq 1$) our error term simplifies to become

$$O(n_{E}) n_{E}^{1/4} - \frac{1}{2} n_{E}^{1/2 - 1/2 a} d_{E}^{1/2 a}$$

(6)

for any $a \geq 1$. Putting $a = 4$ gives the result claimed; it is however clear that there are better choices of $a$. At any rate, the lemma is proved.

We can now state the lower bound result we will need, as

**Corollary.** For a suitable constant $c_{3} > 0$, we have

$$\sum_{N \alpha < x} \frac{1}{N \alpha} \gg \kappa(E) \log x$$

whenever $x \geq c_{3} n_{E}^{d_{E}}$. Moreover, if $E$ has a bad zero $\beta_{E}$ then also

$$\sum_{N \alpha < x} \frac{1}{N \alpha} \gg \frac{\kappa(E)}{1 - \beta_{E}}$$

whenever $x \geq c_{3} n_{E}^{d_{E}}$.

**PROOF:** We have, by Lemma 5,

$$\sum_{N \alpha < x} \frac{1}{N \alpha} \geq \sum_{N \alpha < x} \frac{1}{N \alpha (1 - \frac{N \alpha}{x})^{n_{E}}}$$

$$\geq \kappa(E) \left( \log x - \sum_{j=1}^{n_{E}} \frac{1}{j + \frac{\kappa(O(E))}{\kappa(E)}} - e^{O(n_{E})} n_{E}^{d_{E} 3/8 - 1/2} \right)$$

(7)
Let $c_3 > 0$ be arbitrary for now: then for $x \geq c_3 n_E^3 d_E$ we have
$$n_E d_E^{3/8} \leq n_E^{9/8} d_E^{3/8} \leq c_3^{3/8} n_E^{3/8} x^{3/8},$$
so the last term above is
$$O(n_E) - \frac{3n_E}{8} x^{-1/8} \leq x^{-1/8}$$
provided $c_3$ is large enough.

Moreover by $\sum_{j=1}^{n_E} \frac{1}{j} \leq \log n_E + O(1)$ and by (15) of §2 we see that
$$\sum_{j=1}^{n_E} \frac{1}{j} + \frac{\kappa_0(E)}{\kappa(E)} \geq b_E - \frac{1}{2} \log d_E + \frac{1}{2} n_E (\gamma + \log 2\pi) - \log n_E - O(1)$$
since $r_1(E) + r_2(E) \geq \frac{1}{2} n_E$. The terms involving $n_E$ are bounded below independently of $n_E$ so we have
$$- \sum_{j=1}^{n_E} \frac{1}{j} + \frac{\kappa_0(E)}{\kappa(E)} \geq b_E - \frac{1}{2} \log d_E - C$$
with a suitable constant $C$. Thus (7) becomes
$$\sum_{Na>x} \frac{1}{Na} \geq \kappa(E) (\log x - \frac{1}{2} \log d_E - C + b_E) - x^{-1/8}$$
if $x \geq c_3 n_E^3 d_E$, or more simply (on raising $c_3$ to absorb $C$, if necessary)
$$\sum_{Na>x} \frac{1}{Na} \geq \kappa(E) (\frac{1}{2} \log x + b_E) - x^{-1/8} \quad (8)$$
under the same hypothesis.

Support first that $E$ does not have a bad zero: then
$$\kappa(E) \gg 1/\log d_E,$$
by (2) of §4,
and since $b_E \geq 0$, by (12) of §2, we can deduce from (8) that
\[ \sum_{N \alpha < x} \frac{1}{N \alpha} \geq \kappa(E) \left( \frac{1}{2} \log x - O\left( \frac{\log d_E}{x^{1/8}} \right) \right) \geq \frac{1}{3} \kappa(E) \log x \]

provided \( x \) is large enough, which can be ensured by again raising \( c_3 \) (if necessary). Actually the case \( E = \emptyset \) is not handled by this, because the results of §4 omit it, but this case is immediate from (8) and \( \kappa(\emptyset) = 1 \).

If, on the other hand, \( E \) has a bad zero \( \beta_E \) then by the functional equation \( 1 - \beta_E \) is also a zero (see the paragraph following (1) of §2) and \( \beta_E \) is real so (12) of §2 yields \( b_E \geq 1/1-\beta_E \).

Also by (2) of §4 we have \( \frac{1 - \beta_E}{\kappa(E)} \ll 1 \), so (8) reads

\[
\sum_{N \alpha < x} \frac{1}{N \alpha} \geq \frac{1}{2} \kappa(E) \log x + \frac{\kappa(E)}{1 - \beta_E} \left( 1 - \frac{1 - \beta_E}{\kappa(E)x^{-1/8}} \right) \geq \frac{1}{2} \kappa(E) \log x + \frac{1}{2} \frac{\kappa(E)}{1 - \beta_E}
\]

for large enough \( x \) (ensured by making \( c_3 \) suitably large); but this inequality includes both assertions of the corollary so the proof is complete.

It is to be noted that, if \( \beta_E \) exists and if \( x \) is near the lower bound allowed by the corollary, then \( \frac{\kappa(E)}{1 - \beta_E} \) is significantly larger than \( \kappa(E) \log x \).

§2 An Upper Bound

We now use the average result of Lemma 5 to get an upper bound for the sum of §1: this employs more significant properties of the averaging process, and indeed we will find justification for the term
"average" on the way. The method of proof of lemma 6 is based on the ideas of Landau [17]; we begin with some more general constructions.

We consider only real-valued functions defined on \((0, \infty)\) and call such a function **locally integrable** if it is integrable on \((0, a)\) for every \(a > 0\). If now \(f\) is locally integrable define \(Jf\) by

\[
(Jf)(y) = \int_0^y f(t) \, dt \quad y > 0
\]

and observe that \(Jf\) is continuous on \((0, \infty)\) and that \(\lim_{y \to 0^+} (Jf)(y) = 0\), hence \(Jf\) is also locally integrable and we can form \(J^k f\) for integers \(k \geq 0\).

Of particular interest are summatory functions \(S\) defined as follows: let \(\{a_m\}, m \geq 1\) be a sequence of real numbers and put

\[
S(y) = \sum_{m \leq y} a_m .
\]

Then the term "average" used above is justified by

\[
(J^k S)(y) = \frac{1}{k!} \sum_{m \leq y} a_m (y-m)^k , \quad k \geq 0 \quad (1)
\]

which is proved by induction on \(k\): for \(k = 0\) it is the definition of \(S\) and for \(k \geq 1\) we have

\[
(J^k S)(y) = \int_0^y (J^{k-1} S)(t) \, dt = \int_0^y \frac{1}{(k-1)!} \sum_{m < t} a_m (t-m)^{k-1} \, dt =
\]

\[
\frac{1}{(k-1)!} \sum_{m < y} a_m \int_m^y (t-m)^{k-1} \, dt = \frac{1}{(k-1)!} \sum_{m < y} a_m \frac{(y-m)^k}{k}
\]

as claimed.
To deduce an upper bound for $S(y)$ from knowledge of $J^kS$ for some $k$ we need a relation between them, which is provided by:

if $f$ is locally integrable and $k \geq 1$ then

$$
\int_y^{y+z} \cdots \int_{y_1}^{y_1+z} \int_{y_{k-1}}^{y_{k-1}+z} f(y_k)dy_k \cdots dy_1
$$

$$
= \sum_{\ell=0}^k (-1)^{k-\ell} \binom{k}{\ell} (J^k f)(y+\ell z) \quad \text{for } y,z > 0
$$

and which is again proved by induction on $k$. For $k = 1$ it says only that

$$
\int_y^{y+z} f(y_1)dy_1 = \int_0^{y+z} f(t)dt - \int_0^y f(t)dt
$$

and for $k \geq 2$ we have

$$
\int_y^{y+z} \cdots \int_{y_1}^{y_1+z} f(y_k)dy_k \cdots dy_1 = \int_y^{y+z} \sum_{\ell=0}^{k-1} (-1)^{k-\ell} \binom{k-1}{\ell} (J^{k-1} f)(y+\ell z)dy_1
$$

$$
= \sum_{\ell=0}^{k-1} (-1)^{k-(\ell+1)} \binom{k-1}{\ell} \int_y^{y+(\ell+1)z} (J^{k-1} f)(t)dt = \sum_{\ell=1}^k (-1)^{k-\ell} \binom{k-1}{\ell-1} (J^k f)(y+\ell z)
$$

$$
- \sum_{\ell=0}^{k-1} (-1)^{k-\ell} \binom{k-1}{\ell} (J^k f)(y+\ell z) = \sum_{\ell=0}^k (-1)^{k-\ell} \binom{k}{\ell} (J^k f)(y+\ell z)
$$

by Pascal's triangle.

We can now give an upper bound for a summatory function since

$$
\text{if } a_m \geq 0 \text{ for all } m \text{ then } z^k S(y) \leq \sum_{\ell=0}^k (-1)^{k-\ell} \binom{k}{\ell} (J^k S)(y+\ell z)
$$

for every $k \geq 1$ and $y,z > 0$;

for this we observe that $S(y)$ is increasing hence
As the last step in the preparations we note that we will have to estimate sums as in (3), for which purpose we "invert" the relation (2): namely given $\int k^f$ (without knowing $f$) we must "reconstruct" $f$; then we can apply (2). A simple such result is:

\[
given F \text{ so that } D^k F \text{ is continuous and locally integrable and so that } \lim_{y \to 0^+} D^l F = 0 \text{ for } 0 \leq l \leq k-1; \text{ then} \]

\[
J^k D^k F = F. \tag{4}
\]

Here $D$ is, of course, the differentiation operator. We set

\[
G = J^k D^k f \quad \text{and begin the proof by observing}
\]

\[
J^l D^k F = D^{k-l} G, \quad 0 \leq l \leq k \tag{5}
\]

because $J^l D^k F$ is continuous and locally integrable implies that $D^{k-l} J^{k-l}$ acts as the identity on it hence $J^l D^k F = D^{k-l} J^{k-l} J^l D^k F = D^{k-l} J^k D^k F = D^{k-l} G$, as claimed. We now conclude the proof by showing that $D^{k-l} G = D^{k-l} F$ for $0 \leq l \leq k$ by induction on $l$: for $l = 0$ this is contained in (5). Moreover if $1 \leq l \leq k$ and $D^{k-(l-1)} G = D^{k-(l-1)} F$ then $D D^{k-l} G = D D^{k-l} F$ so $D^{k-l} G - D^{k-l} F$ is a constant $C$; but as $y \to 0^+$, $D^{k-l} G \to 0$ by (5) and $D^{k-l} F \to 0$ by hypothesis, hence $C = 0$ and $D^{k-l} G = D^{k-l} F$ completing the induction.

We are now prepared for
LEMMA 6. We have, for \( y \leq 1 \),

\[
\sum_{N(a) < y} \frac{1}{N(a)} \leq \kappa(E) \log y + (\kappa_0(E) + \kappa(E) \log 2) + \mathcal{O}(e^{-\frac{3}{8} n_E d_E y^{-1/2}})
\]

Proof: We apply the above method with

\[ k = n_E \quad \text{and} \quad z = y/n_E \]

and write \( H = \sum_{j=1}^{k} 1/j \). Now if

\[ a_m = \sum_{N(a) = m} \frac{1}{N(a)} \geq 0 \]

then

\[ S(y) = \sum_{N(a) < y} \frac{1}{N(a)} \]

is the sum we must estimate, and by (1) we have

\[ (J^k S)(y) = \frac{y^k}{k!} \sum_{N(a) < y} \frac{1}{N(a)} (1 - \frac{N(a)}{y})^k \]

to which lemma 5 applies. Putting all of this into the inequality (3) results in

\[ z^k \sum_{N(a) < y} \frac{1}{N(a)} \leq \kappa(E) \sum_A + (\kappa_0(E) - \kappa(E)) \sum_B + \sum_C \]

(6)

by lemma 5, where
\[
\sum_A = \sum_{\ell=0}^{k} (-1)^{k-\ell} \binom{k}{\ell} \frac{(y+\ell z)^k}{k!} \log(y+\ell z)
\]

\[
\sum_B = \sum_{\ell=0}^{k} (-1)^{k-\ell} \binom{k}{\ell} \frac{(y+\ell z)^k}{k!}
\]

\[
\sum_C = \sum_{\ell=0}^{k} (-1)^{k-\ell} \binom{k}{\ell} \frac{(y+\ell z)^k}{k!} 0(e^{O(k)} k d_E^{3/8} (y+\ell z)^{-1/2})
\]

are to be estimated. We begin with the last and in

\[
\sum_C \ll e^{O(k)} k d_E^{3/8} \sum_{\ell=0}^{k} \binom{k}{\ell} \frac{(y+\ell z)^{k-1/2}}{k!}
\]

we estimate the sum trivially as

\[
\frac{(y+\ell z)^{k-1/2}}{k!} \sum_{\ell=0}^{k} \binom{k}{\ell} = \frac{(2kz)^k (2y)^{-1/2}}{k!} 2^k
\]

\[
\ll 4 k \frac{k^k}{k!} \frac{k-1/2}{k} \leq 4 k e z y
\]

where we have used \(2y = y + kz = 2kz\). Thus

\[
\sum_C \ll e^{O(k)} k d_E^{3/8} z y
\]  \hspace{1cm} (7)

and we turn to \(\sum_B\): putting \(F(y) = \frac{y^k}{k!}, y > 0\) it is easy to verify the hypothesis of (4), hence by \((D^k F)(y) = 1\) and (2) we get

\[
\sum_B = \int_y^{y+z} \int_y^{y+2z} \cdots \int_y^{y+kz} 1 dy_{k-1} \cdots dy_1 = z^k
\]  \hspace{1cm} (8)

and we can finally turn to the more difficult term \(\sum_A\): proceeding as
above, we put
\[ F(y) = \frac{y^k}{k!} \log y, \quad y > 0 \]

and define \( c_j = \sum_{j=k+1}^{k} \frac{1}{\xi} \) for \( 0 \leq j \leq k \).

By induction it is easily verified that
\[ (D^j F)(y) = \frac{y^{k-j}}{(k-j)!} (\log y + c_j) \quad \text{for} \quad 0 \leq j \leq k \]
hence the hypothesis of (4) are satisfied and noting that \( c_k = H \) the relation (2) gives us
\[ \sum_A = \int y^{y+z} \int y_1^{y+z} \cdots \int y_{k-1}^{y+z} (\log y_k + H) dy_k \cdots dy_1 \]
\[ \leq \log(y^k + H) \int y^{y+z} \cdots \int y_{k-1}^{y+z} 1 dy_k \cdots dy_1 = z^k (\log 2y + H) \]
since \( \log y + H \) is increasing. Putting this with (7), (8) in (6) gives finally
\[ \sum \frac{1}{Na} \leq \kappa(E)(\log y + \log 2 + H) + (\kappa_0(E) - H \kappa(E)) + O(e^{0(k) k d_E y^{-1/2}}) \]
completing the proof.
COROLLARY. Let $B > 0$; then for $x \geq 1$,

$$\sum_{x \leq N(a) < x} \frac{1}{N} \leq B \kappa(E) \log x + \left( \sum_{j=1}^{n_E} \frac{1}{x} + \log 2 \right) \kappa(E) + O(e^{-n_E n_E d_E 3/8} x^{-1/2})$$

Proof: Write $\sum = \sum_{N a < x}^{B+1} - \sum_{N a < x}^{B+1}$. Then lemma 6 supplies an upper bound for $\sum_{N a < x}^{B+1}$, and an upper bound for $- \sum_{N a < x}^{B+1}$ follows from the lower bound (7) of §1 for $\sum_{N a < x}^{B+1}$. It is particularly important that the $\kappa_0(E)$ terms cancel; the significance of this is clear from the corollary to lemma 5.

§3 An analogous sum for prime ideals

The effect of the existence of a bad zero $\beta_E$ for $E$ on the sum $\sum_{N a < x}^{B+1} 1/Na$ is to force it to be large (corollary to lemma 5); similarly the effect of the existence of $\beta_E$ on the sum $\sum_{x \leq Np < x}^{B} 1/Np$ over prime ideals is to force it to be small, which is seen by comparing the next result with the inequality (19) of §6. The proof follows the idea of Bombieri [2] referred to in §1 and is now quite easy since we have established the main results in the previous sections.
**THEOREM 2.** There is a constant $c_4 > 0$ so that if $E$ has a bad zero $\beta_E$, and if $B \geq 1$, then

$$\sum_{x \leq Np < x^B} 1/Np \ll B^2 (1-\beta_E) \log x$$

for every $x \geq c_4 n_E d^2 E$.

**Proof:** Let $x$ satisfy the above condition where $c_4$ is yet to be determined. Consider the product

$$\prod_{N(a) < x} \frac{1}{N(a)} \prod_{x \leq Np < x^B} \frac{1}{N(p)} = \sum_{x \leq Np < x^B} \frac{1}{N(ap)}$$

Then on the right hand side we have a sum over certain integral ideals $B$ of the form $B = ap$ where $a,p$ satisfy the stated conditions; certainly $x \leq N(B) < x^{B+1}$ holds for any such $B$. Conversely, given $B$ with $x \leq N(B) < x^{B+1}$ we can express $B$ in the form $ap$ in $< B + 1$ ways: for if $p_1, \ldots, p_v$ are the (necessarily different) prime ideals which participate in such a representation then $p_1 \ldots p_v$ divides $B$ hence $x^v \leq N(p_1) \ldots N(p_v) \leq N(B) < x^{B+1}$. Thus

$$\sum_{N(a) < x} \frac{1}{Na} \sum_{x \leq Np < x^B} \frac{1}{Na} \ll B \sum_{x \leq N(a) \leq x^{B+1}} \frac{1}{Na}$$

since $B + 1 \leq 2B$. Applying the corollary to lemma 5 (when $\beta_E$ exists) and the corollary to lemma 6 shows
for a certain constant $C$, which yields

\[
\sum_{x \leq NP < B} 1/NP \ll B^2 \left( \kappa(E) \log x + \log Cn \right)
\]

since the condition on $x$ allows absorption of the $\log Cn$ term for large enough $c_4$.

It remains to show that the $0$-term remains bounded for large enough $c_4$, so we need an upper bound for $1/\kappa(E)$ for which we apply theorem 1' (combined with (2) of §4):

\[
1/\kappa(E) \ll \max(n_1 \log d, d^{1/2}) \leq \left( n_1 \log d \right)^{1/2}
\]

by $n_1 \geq 2$, since $\beta Q$ does not exist.

Since $\frac{\log d}{\log x} \leq 1$ our error term (1) becomes

\[
\sum_{x \leq NP < B} 1/NP \ll B^2 \left( \kappa(E) \log x + \log Cn \right)
\]

which is indeed small for suitable $c_4$, completing the proof.
Actually the lower bound (2) for \( \kappa(E) \) is an extremely weak application of theorem 1', the significance of which depends, precisely on the exponent \( 1/n_E \) of \( d_E \), and the full strength of which will be needed later. There are also lower bounds for \( \kappa(E) \) which come from the geometry of numbers: for \( \kappa(E) = \frac{r_1(E)}{2} \frac{r_2(E)}{(2\pi)} R_E h_E \) (from p.161 of Lang [18], for example). Using the trivial estimates \( w_E \ll n_E^2 \) and \( h_E \gg 1 \) this becomes

\[
\kappa(E) \geq e^{\frac{0(n_E)}{d_E - 1/2} R_E}
\]

which, together with the lower bound of Remak [24]:

\[
R_E \geq e^{\frac{0(n_E)}{n_E}}
\]

for the regulator, gives a slightly better lower bound for \( \kappa(E) \) than (2), thus avoiding theorem 1'.

The reason for pointing out this alternative is that it is only this occurrence of \( n_E \) (namely, from theorem 1') which does not arise from the use of averaging processes (as mentioned in the Introduction). In addition to the situations mentioned in \( \S 4 \) when \( n_E \) can be avoided we also want to observe that it is known that (see Pohst [22])

\[
R_E \gg 1 \quad \text{if } E \text{ is totally real}
\]

which again avoids the additional \( n_E \). Unfortunately when we return to the exceptional zero terminology it is difficult to appreciate the restrictions imposed by these alternatives e.g. (5) applies only when
$K$ is totally real, and we restrict $\chi$ by insisting it be unramified at all the archimedean valuations of $K$. 
CHAPTER III

MEAN VALUES OF DIRICHLET POLYNOMIALS

The single theme of the present chapter is to obtain upper bounds for the mean value

$$\int_{-T}^{T} \left| \sum_{m} b_m e^{it} \right|^2 dt$$

of the title in the form of averaged partial sums of the coefficients $b_m$. This is the technical crux of the present approach. The point is that a suitable form of averaging allows estimates which are far sharper than what is possible before averaging, a fact which is particularly important as the degree of the number field $K$ becomes large: indeed the average we use will depend on the degree of $K$, by virtue of the rate of growth of $L$-series (defined on $K$) in the critical strip. Moreover, and this is equally important, the method is compatible with the application of the Selberg sieve.

§1 Averaging functions and their properties

Let $\eta$ be a bounded, integrable function on $\mathbb{R}$ (in the sense of Lebesgue measure), normalize its Fourier transform as

$$\hat{\eta}(t) = \int_{\mathbb{R}} \eta(x) e^{-itx} dx \quad (t \in \mathbb{R})$$

and define its'"multiplicative analogue" by

$$H(y) = \eta(\log y) \quad (y \in \mathbb{R}, y > 0)$$

We call $\eta$ an averaging function (for motivation) and begin with the formal
LEMMA 7. If \( \eta \) is an averaging function and \( \{b_m\}_{m \geq 1} \) is any sequence of complex coefficients so that \( \sum_{m} |b_m| < \infty \) then

\[
\frac{1}{2\pi} \int_{\mathbb{R}} |\hat{\eta}(t)|^2 \left| \sum_{m} b_m e^{-it} \right|^2 dt = \int_{0}^{\infty} \left| \sum_{m} b_m \eta(y/m) \right|^2 \frac{dy}{y}
\]

PROOF: Define the sequence \( \mu_m = \log m \) (\( m \geq 1 \)) and put

\[
F(t) = \sum_{m} b_m e^{-i\mu_m t}
\]

a bounded integrable function since

\[
\sum_{m} |b_m| < \infty.
\]

Put

\[
C(x) = \sum_{m} b_m \eta(x-\mu_m)
\]

and

\[
C_*(x) = \sum_{m} |b_m| |\eta(x-\mu_m)|
\]

then \( C_* \) is bounded and measurable, hence, by the monotone convergence theorem and the translation invariance of Lebesgue measure, we have

\[
\int_{\mathbb{R}} C_*(x) \, dx = \sum_{m} |b_m| \int_{\mathbb{R}} |\eta(x-\mu_m)| \, dx = \left( \sum_{m} |b_m| \right) \int_{\mathbb{R}} |\eta(x)| \, dx < \infty,
\]

so \( C_* \) is integrable. Since \( C \) is then also bounded and integrable it follows that \( C \) is square integrable.

Moreover since \( C_* \) is integrable we have

\[
\hat{C}(t) = \int_{\mathbb{R}} \sum_{m} b_m \eta(x-\mu_m) e^{-ixt} \, dx = \sum_{m} b_m e^{-i\mu_m t} \int_{\mathbb{R}} \eta(x-\mu_m) \, dx = F(t) \hat{\eta}(t),
\]

by the dominated convergence theorem and the translation invariance of Lebesgue measure again.

But \( C \) is square integrable so, with our normalization of Fourier
transforms, Plancherel's theorem gives

\[ \frac{1}{2\pi} \int_{\mathbb{R}} |\hat{C}(t)|^2 dt = \int_{\mathbb{R}} |C(x)|^2 dx \]

Here the left hand side is what was claimed and the right hand side, on substituting \( x = \log y \), becomes

\[ \int_{0}^{\infty} |C(\log y)|^2 \frac{dy}{y} = \int_{0}^{\infty} |\sum_{m} b_{m} \eta(\log y - \log m)|^2 \frac{dy}{y}, \]

again as claimed.

This is simply a formal generalization of theorem 1 of Gallagher [10], and, as there, is to be applied in the form

\[ \frac{1}{2\pi} \min_{t \in [-T,T]} \left( |\hat{\eta}(t)|^2 \right) \int_{-T}^{T} \left| \sum_{m} b_{m} \right|^2 dt \leq \int_{0}^{\infty} \left| \sum_{m} b_{m} H(y/m) \right|^2 \frac{dy}{y} \]  

so any application is dependent on finding an averaging function \( \eta \) so that \( \hat{\eta} \) is not "too small" on \([-T,T]\) and so that the sums \( \sum_{m} b_{m} H(y/m) \) can be estimated "effectively". The basis of the estimation of such sums will be the relationship between the Fourier and the Mellin transforms: that is, to the averaging function \( \eta \) we associate the function

\[ h(s) = \int_{-\infty}^{\infty} \eta(x)e^{-xs} dx \]

where now \( s \) is a complex variable. Clearly

\[ \hat{\eta}(t) = h(it) \quad (t \in \mathbb{R}) \]
but the summation method depends on $h$ being analytic in a vertical strip, hence on $\eta$ being more special than was initially assumed. Under fairly general conditions (which will not be needed for our choice of $\eta$), corresponding to (4) there is the inversion formula for the Mellin transform

$$H(y) = \frac{1}{2\pi i} \oint_{c-i\infty}^{c+i\infty} h(s)y^s ds \quad \text{(for suitable $c \in \mathbb{R}$)}$$

(6)

which will be used to estimate sums of the form $\sum h_m H(y/m)$.

We now construct a sequence $\eta_\ell$ ($\ell$ a positive integer) of special averaging functions and let $H_\ell, h_\ell$ be the functions associated to it by (2) and (4). Namely fix a parameter $\Lambda > 0$ and put

$$\eta_1(x) = \begin{cases} \frac{\Lambda}{2}, & |x| < 1/\Lambda \\ \frac{\Lambda}{4}, & |x| = \frac{1}{\Lambda} \\ 0, & |x| > 1/\Lambda \end{cases}$$

(7)

the function of Gallagher's theorem 1 mentioned earlier. Defining, as usual, the convolution of functions $f, g$ by

$$(f * g)(x) = \int f(x-u)g(u)du$$

(8)

we get our sequence $\eta_\ell$ by defining inductively

$$\eta_\ell = \eta_1 * \eta_{\ell-1} \quad (\ell \geq 2).$$

(9)

As examples of how the functions $\eta_\ell$ look, let us record:
\[ n_2(x) = \begin{cases} \frac{A}{2}(1 - \frac{1}{2}|x|), & |x| \leq 2/A \\ 0, & |x| \geq 2/A \end{cases} \]

\[ n_3(x) = \begin{cases} \frac{3A}{4} \left( \frac{A}{4} x^2 \right), & |x| \leq 1/A \\ \frac{9A}{16}(1 - \frac{1}{3}|x|)^2, & \frac{1}{A} \leq |x| \leq \frac{3}{A} \\ 0, & |x| \geq 3/A \end{cases} \]

which already exhibits the tendency of the \( n_\ell \) to become smoother and "more spread out" as \( \ell \) increases.

We require only some simple properties of the \( n_\ell \):

\( n_\ell \) is a non-negative even function vanishing outside the interval \((-\ell/A, \ell/A)\) and satisfying

\[ n_\ell(x) \leq \frac{A}{2} \quad (x \in \mathbb{R}) \quad (10) \]

For these properties are clear for \( \ell = 1 \) and are then readily verified from (9) by induction. For example, from \( n_{\ell-1}(x) \leq \frac{A}{2} \) follows

\[ n_\ell(x) = \int_{-\ell/A}^{\ell/A} \frac{A}{2} n_{\ell-1}(x-u) du \leq \frac{A^2}{4} \int_{-1/A}^{1/A} du = \frac{A}{2}. \]

Since \( n_\ell \) vanishes outside a bounded set the integral (4) converges for all complex \( s \), and indeed

\[ h_\ell(s) = \left( \frac{\sinh(A^{-1}s)}{A^{-1}s} \right)^\ell \quad \text{is entire.} \quad (11) \]

For \( h_1(s) = \int_{-1/A}^{1/A} \frac{A}{2} e^{-xs} dx = \frac{A}{2s} \left( e^{s/A} - e^{-s/A} \right) \), as required, and, by (9)
we have 

\[ h_{\lambda}(s) = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} n_1(u) n_{\lambda-1}(x-u) du \right) e^{-x^2} dx \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} n_1(u) e^{-us} n_{\lambda-1}(x-u) e^{-(x-u)^2} du dx \]

\[ = \int_{-\infty}^{\infty} n_1(u) e^{-us} \int_{-\infty}^{\infty} n_{\lambda-1}(x-u) e^{-(x-u)^2} dx du = h_{\lambda-1}(s) \int_{-\infty}^{\infty} n_1(u) e^{-us} du \]

\[ = h_{\lambda-1}(s) h_1(s), \text{ hence (11) by induction.} \]

We next estimate growth of \( h_{\lambda}(s) \): from \( |\sinh(s)| = \left| \frac{e^s - e^{-s}}{2} \right| \)

\[ \leq e^\sigma + e^{-\sigma} \leq e |\sigma| \]

follows

\[ |h_{\lambda}(s)| \leq \frac{\Lambda^2 e^{-1/2} |\sigma|}{|s|^2} \]  \( (12) \)

which shows that \( h_{\lambda}(s) \) is small when \( s \) is large, whenever \( s \) is constrained to lie in a strip \( \sigma_0 \leq \sigma \leq \sigma_1 \). However, this estimate hardly reflects the analyticity of \( h_{\lambda}(s) \) at \( s = 0 \), for which we need

\[ |h_1(s) - 1| \leq \frac{|s|^2}{5\Lambda^2} \quad \text{for} \quad |s| \leq \Lambda, \]  \( (13) \)

for which it suffices to consider the Taylor series of \( \frac{\sinh(s)}{s} - 1 \) in \( |s| \leq 1 \): we have

\[ \left| \frac{\sinh(s)}{s} - 1 \right| \leq \sum_{k=1}^{\infty} \frac{|s|^{2k}}{(2k+1)!} \]

\[ \leq |s|^2 \sum_{k=1}^{\infty} \frac{1}{(2k+1)!} = |s|^2 (\sinh 1 - 1) \leq \frac{|s|^2}{5} \]

as required. Of course:
Finally we proceed to the justification of (6) in our special case: the inversion formula for the Fourier transform gives

\[ \eta_{\frac{1}{2}}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \eta(t)e^{ixt} dt \quad x \in \mathbb{R} \quad (14) \]

in our normalization, where the integral is improper for \( \ell = 1 \) but converges absolutely for \( \ell \geq 2 \) by (12), in view of (5). This follows from the simplest forms of the inversion theorem or can be verified directly: for \( \ell = 1,2 \) it follows readily from the familiar integral

\[ \int_{0}^{\infty} \frac{\sin \alpha t}{t} dt = \begin{cases} \frac{\pi}{2}, & \alpha > 0 \\ 0, & \alpha = 0 \\ -\frac{\pi}{2}, & \alpha < 0 \end{cases} \]

and then, for \( \ell \geq 3 \), we can use induction, because of (9), and because the integral (14) converges absolutely when \( \ell - 1 \geq 2 \). Now by (2) and (5) we have

\[ H_{\frac{1}{2}}(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} h_{\frac{1}{2}}(it)e^{it\log y} dt = \frac{1}{2\pi i} \int_{-\infty}^{\infty} h_{\frac{1}{2}}(s)y^{s} ds , \quad y > 0 \]

which is (6) with \( c = 0 \), and we can extend this to all real \( c \) as follows: suppose \( c > 0 \) (the case \( c < 0 \) is similar) and consider the positively oriented rectangle with vertices \(-iU, c - iU, c + iU, iU\) where \( U > 0 \) is large. Since \( |\sigma| \leq c \) on our rectangle it follows, from (12), that the integral of \( h_{\frac{1}{2}}(s)y^{s} \) along the top and bottom edges is

\[ \leq \Lambda e^{\Lambda \frac{1}{2}c} y^{c} \int_{0}^{c} \frac{1}{|\sigma + iU|^{\frac{1}{2}}} \frac{d\sigma}{2\pi} \ll \frac{1}{U}, \text{ which tends to zero as} \]

(13) applies to \( h_{\frac{1}{2}}(s) \) for all \( \ell \), via \( h_{\frac{1}{2}}(s) = h_{\frac{1}{2}}(s)^{\ell} \).
\( U \to \infty \). Thus by Cauchy's theorem we get

\[
\lim_{U \to \infty} \frac{1}{2\pi i} \int_{-iU}^{iU} h_\xi(s) y^s ds = \lim_{U \to \infty} \frac{1}{2\pi i} \int_{c-iU}^{c+iU} h_\xi(s) y^s ds ,
\]

and (6) is justified for any real \( c \).

We summarize the basic features of the method in

**Lemma 7'**. For the functions \( \eta_\xi \), defined by (7) and (9), we have

\[
H_\xi(y) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} h_\xi(s) y^s ds , \quad \text{for any real } c.
\]

Moreover if the parameter \( \Lambda \) satisfies \( \Lambda \geq \xi^{1/2} T \) then for any sequence \( \{b_m\}_{m \geq 1} \) of complex numbers satisfying \( \sum |b_m| < \infty \) we have

\[
\sum_{m} |b_m| \int_{-T}^{T} \frac{|H_\xi(y/m) - H_\xi(y)|^2}{y} \, dt < K
\]

where the implied constant is absolute.

**Proof**: By (3) it suffices to show \( |\hat{\eta}_\xi(t)| \gg 1 \) for \( |t| \leq T \). By (5),(13) and \( h_\xi(s) = h_1(s)^\xi \) we have, for \( |t| \leq T \),

\[
|\hat{\eta}_\xi(t)| \geq \left(1 - \frac{|t|^2}{5A^2}\right)^\xi \geq \left(1 - \frac{T^2}{5A^2}\right)^\xi \geq \left(1 - \frac{1}{5\xi}\right)^\xi \gg 1 .
\]

In chapter V we will need to know that:

\[
|h_\xi(s)| \quad \text{is bounded (independent of } \xi \text{ and } \Lambda) \quad (15)
\]

in the strip \( |\sigma| \leq \xi^{-1/2} \Lambda \).
To see this let $z = x + iy$ be a complex number; we claim:

$$\left| \frac{\sinh z}{z} \right| \leq \frac{\sinh |x|}{|x|}$$

Now $\sinh z = \sinh x \cos y + i \cosh x \sin y$ hence we have $|\sinh z|^2 = \sinh^2 x + \sin^2 y$, so our claim is equivalent to:

$$\frac{\sinh^2 x + \sin^2 y}{x^2 + y^2} \leq \frac{\sinh^2 x}{x^2}.$$ But $\left| \frac{\sin y}{y} \right| \leq 1 \leq \frac{\sinh x}{x}$ makes this clear, and settles the claim. It is now easy to prove (15) for, by the claim and then (13), we have

$$|h_\ell(s)| \leq (h_1|\sigma|)^\ell \leq (1 + \frac{|\sigma|^2}{5A^2})^\ell \leq (1 + \frac{1}{5})^\ell \ll 1.$$ 

Finally, we remark that it is by no means clear that these functions are optimal in any sense: they are however sufficient for the applications.

§2 An inequality of Polya-Vinogradov type, for ideals.

We estimate some special auxilliary sums of the form $\sum_{m} b_m H_\ell(y/m)$ depending on our field $K$; more precisely, we fix $K$ again and consider Dirichlet polynomials $\sum_{a} b(a)N(a)^{-s}$ (where $a$ runs through the integral ideals of $K$) in place of the more general Dirichlet polynomials of §1. This shift of emphasis is more than just a change of notation because, when we apply Selberg's sieve method, we need it to be Selberg's sieve for the field $K$, not for $\mathbb{Q}$. 
First some notation: define arithmetic functions $\phi = \phi_K$ (Euler $\phi$-function) and $\omega = \omega_K$ on the integral ideals of $K$ by

$$\phi(a) = N(a) \prod_{p/a} \left(1 - \frac{1}{N(p)}\right)$$

$$\omega(a) = \text{number of prime divisors of } a$$

not counting multiplicities

Moreover we must now consider $L$-series for (possibly) imprimitive characters i.e. let $\chi \mod m$ be a congruence class character of $K$ with conductor $f_\chi$, so $\chi$ is induced by the primitive character $\chi^* \mod f_\chi$, and put

$$D_\chi = 4^{\omega(m) - \omega(f_\chi)} d_{\chi^*}$$

$$\delta_\chi = \delta_{\chi^*} = \begin{cases} 1 & , \chi^* = 1 \\ 0 & , \text{otherwise} \end{cases}$$

Then, using the averaging functions $\mu_\lambda$ of §1 based on the parameter $\lambda > 0$, we can state

**Lemma 8.** If $\chi \mod m$ is a (possibly imprimitive) character, and if

$A \gg 1$, then for $y > 0$ we have

$$|\sum_{a} \frac{\chi(a)}{N(a)} \prod_{n} (y/N(a)) - \delta_\chi \phi(m) e(n \lambda) n^{-1} \frac{1}{2^{n+1}} \frac{1}{n!} \frac{1}{\log \lambda} n^{-1}|$$

**Proof:** Let $a \geq 2$ be chosen later. We will actually prove somewhat more: namely, the value $\lambda = 2n$ in the statement of the lemma is only one of many possibilities. Indeed, we show that the above statement
holds for any \( \ell \) satisfying \( \ell \geq 2 \), \( \ell \geq \frac{n}{2}(1+\frac{1}{a}) + \frac{5}{4} \), and \( \ell \ll n \).

In particular \( \ell = 2n \) satisfies these conditions no matter what value we assign to \( a \geq 2 \).

Now by §1 we have for \( y > 0 \):

\[
H_\ell(y/Na) = \frac{1}{2\pi i} \int_{\frac{1-i\infty}{a}}^{\frac{1+i\infty}{a}} h_\ell(s) \frac{y^s}{N(a)_s} ds
\]

where the integral converges absolutely, by \( \ell \geq 2 \). Moreover the Dirichlet series \( L(s+\chi, x) = \sum a N(a)^{s+1} \) converges absolutely uniformly on the line \( \sigma = \frac{1}{a} \), hence

\[
\sum_{a N(a)}^{\chi} H_\ell(y/Na) = \frac{1}{2\pi i} \int_{\frac{1-i\infty}{a}}^{\frac{1+i\infty}{a}} L(s+\chi, x) h_\ell(s) y^s ds .
\]

We propose moving the line of integration from \( \sigma = \frac{1}{a} \) to \( \sigma = -1 \), and for this purpose consider the positively oriented rectangle \( R_U \) with vertices \( \frac{1}{a} - iU, \frac{1}{a} + iU, -1 + iU, -1 - iU \) where \( U > 0 \) is large (we let \( U \to \infty \)). Note that, for any \( U > 0 \), the function \( L(s+\chi, x) h_\ell(s) y^s \) is analytic inside \( R_U \), with the exception of a simple pole at \( s = 0 \) in case \( \chi = 1 \): in this case we have

\[
L(s, \chi) = \left[ \text{H}(1-\frac{1}{\rho^{1/m}}) N(\rho)^s \right] K(s)
\]

hence the residue of \( L(s+\chi, x) h_\ell(s) y^s \) at \( s = 0 \) is
\[ \prod_{p/m} \left( 1 - \frac{1}{Np} \right)^{\chi(K)} h_{\chi} \left( 0 \right) y^0 = \frac{\varphi(m)}{N(m)} \chi(K). \]

We now show that the integral of \( L(s+1, \chi) h_{\chi}(s)y^s \) along the horizontal edges of \( R_u \) tends to 0 as \( U \to \infty \): for this purpose we can estimate qualitatively. Now

\[ L(s, \chi) = \prod_{p/m, p \not| \chi} \left( 1 - \chi(p) N(p)^{-s} \right) L(s, \chi^*) \quad (4) \]

from (15), hence by lemma 1 we have

\[ L(s+1, \chi) \ll L(s+1, \chi^*) \ll e^{-\frac{1}{2}Np} \left( \frac{1}{2} - \sigma \right) \frac{n(\frac{1}{2} - \sigma)}{\chi} \left( 1 + |t| \right)^{-\frac{1}{2} \chi} \ll U^{-\frac{1}{2}} \]

on the horizontal edges of \( R_u \), so since, by (12) of III\$1,

\[ \left| h_{\chi}(s)y^s \right| \leq \frac{\Lambda}{|s|}e^{-\Lambda s} |\sigma| \max(y^{-1}, y^{1/a}) \ll U^{-\frac{1}{2}} \]

also holds there, and since these edges have length \( 1 + \frac{1}{a} \ll 1 \), the integral of \( L(s+1, \chi) h_{\chi}(s)y^s \) is indeed \( \ll U^{n/2}(1+1/a)U^{-\frac{1}{2}} \ll U^{-5/4} \), which tends to 0 as \( U \to \infty \). This, together with (3), shows that

\[ \frac{\chi(a)}{a N(a)} H_{\chi}(y/Na) = \delta \chi_{\chi N(m)} \kappa(K) - \frac{1}{2\pi i} \int_{-1-i\infty}^{-1+i\infty} L(s+1, \chi) h_{\chi}(s)y^s ds \quad (5) \]

in view of the calculation of the residue in case \( \delta = 1 \). It therefore remains only to estimate the integral remaining in (5), for which purpose we first estimate \( L(s+1, \chi) \) on \( \sigma = -1 \) in the manner employed above: first by (4) we have on \( \sigma = -1 \)

\[ \left| L(s+1, \chi) \right| \leq \prod_{p/m, p \not| \chi} \left( 1 + |N(p)|^{-s-1} \right) |L(s+1, \chi^*)| \leq 2^{\omega(m) - \omega(h_{\chi})} |L(s+1, \chi^*)| \]
and to $L(s+l,x^*)$ we can apply Lemma 1 and get (on $\sigma = -1$)

$$|L(s+l,x^*)| \leq e^{O(n)\frac{1}{a}d\frac{n}{2}(1+\frac{1}{a})} (1+|t|)^{\frac{n}{2}(1+\frac{1}{a})}$$

since the $\left(\frac{s}{s+3}\right)^{\delta}$ term of Lemma 1 is uniformly bounded below on $\sigma = -1$. Thus our integral is

$$\ll 2^{\omega(m)-\omega(h_{\chi})} e^{O(n)\frac{1}{a}d\frac{n}{2}(1/2a - 1)} \int_{-1-\infty}^{-1+\infty} (1+|t|)^{\frac{n}{2}(1+\frac{1}{a})} |h_{\chi}(s)| |ds|$$

$$\ll e^{O(n)\frac{1}{a}d\frac{n}{2}(1/2a - 1)} \int_{0}^{\infty} (1+|t|)^{\frac{n}{2}} |h_{\chi}(-1+it)| dt ,$$

since the integrand is an even function of $t$. To estimate this remaining integral, we break it into two parts so allowing use of the estimates (21) and (13) of §1: supposing $A \geq 2$ (say), then

$$|t| \leq (A^2-1)^{1/2}$$

implies $|-1+it| \leq A$ hence, by (13) of §1, we have

$$|h_{\chi}(-1+it)| \leq (\frac{6}{3})^{\frac{1}{2}} e^{O(n)} \chi_{\mathbb{A}} \ll n ,$$

so we get

$$(A^2-1)^{1/2} \int_{0}^{\infty} (1+|t|)^{\frac{n}{2}(1+\frac{1}{a})} |h_{\chi}(-1+it)| dt \leq e^{O(n)\frac{n}{2}(1+\frac{1}{a})} + 1.$$}

On the other hand, we use (12) of §1 to get

$$\int_{(A^2-1)^{1/2}(1+|t|)^{\frac{n}{2}(1+\frac{1}{a})}}^{\infty} |h_{\chi}(-1+it)| dt \leq \int_{A/2}^{\infty} A^{\frac{\ell}{2}/A} \frac{\ell}{2} (1+|t|)^{\frac{n}{2}(1+\frac{1}{a})} dt$$

$$\leq e^{O(n)A^{\ell}} \int_{A/2}^{\infty} |t|^\frac{n}{2}(1+\frac{1}{a}) - \ell dt = e^{O(n)A^{\ell}} \frac{\frac{n}{2}(1+\frac{1}{a})+1-\ell}{|\frac{n}{2}(1+\frac{1}{a})+1-\ell|} \ll e^{O(n)A^{\ell}}.$$
since \( \frac{n}{2} \left( 1 + \frac{1}{a} \right) + 1 - \varepsilon \leq -\frac{1}{4} \) implies the convergence of the last integral and also that the resulting denominator is bounded below, independent of all the variables. In view of these estimates (6) becomes

\[
\ll e^{\Theta(n)} \frac{1}{a} D^{1/2} a n^{1/2a} y^{-1} A^{\frac{n}{2} a (1 - a) + 1} = e^{\Theta(n)} \frac{1}{a} D^{1/2} A^{\frac{n}{2} a + 1} y^{-1} (a D^{1/2} A^{1/2})
\]

and it remains only to choose \( a \geq 2 \) appropriately. But when

\[
a = \frac{\log D A^{n}}{2n}
\]

then \( a D^{1/2} A^{n/2a} = e^{\Theta(n)} \left( \frac{\log D A^{n}}{2n} \right)^n = \Theta(n) \left( \log D A^{n} \right)^n
\]

which is the value claimed (and is, of course, the minimum value of that function). Finally since \( D \geq 1 \) we have

\[
a \geq \frac{\log A}{2} \geq 2, \text{ provided } A \gg 1,
\]

and the lemma is proved.

From the proof of the lemma it is clear that \( A \) can be taken quite small and, indeed, if we replace the use of \( D \geq 1 \) by the Minkowski lower bound for discriminants then even smaller values of \( A \) are admissible, but we shall not need this. What is critical is that in the error term of the lemma we have \( D^{1/n} \) to a constant power (independent of \( K \)), so that when \( A \) is about \( D \) then the error term becomes small when \( y \) is \( D \) to a constant power.

By a congruence class group \( H \mod m \) we understand a subgroup \( H \) of \( I(m) \) which contains \( P_m \); or, equivalently a subgroup of \( I(m)/P_m \).

As an analytical measure of the "size" of \( H \) we use
\[ h_H = (I(m):H) \]  \hspace{1cm} (8)

\[ D_H = \max\{D : \chi(H) = 1\} \]

where the condition \( \chi(H) = 1 \) is understood as meaning that the character \( \chi \mod m \) is trivial on \( H \).

**Corollary.** Let \( C \) be a coset of the congruence class group \( H \mod m \) and let \( n \) be an integral ideal prime to \( m \). Then

\[
\left| \sum_{a \in C} \frac{1}{\mathcal{N}(a)} \mathcal{H}_2(n) \left( \frac{\chi}{\mathcal{N}(a)} \right) - \frac{1}{\mathcal{N}(n)} \frac{\phi(m)}{\mathcal{N}(m)} \frac{\kappa(K)}{h_H} \right| \ll \mathcal{O}(n) \frac{1}{D_H A} \frac{1}{2^{n+1}} \left( \log \frac{D_H n}{A} \right)^n \]

provided \( A \gg 1 \).

**Proof:** Let \( C_1 \) be the coset of \( H \) containing \( n \) and \( C' = C_1 \backslash C \): writing \( a \) in the sum as \( a = nB \), then \( B \) is an integral ideal in \( C' \) and every \( a \) arises in this way, hence

\[
\sum_{a \in C} \frac{1}{\mathcal{N}(a)} \mathcal{H}_2(n) \left( \frac{\chi}{\mathcal{N}(a)} \right) = \sum_{B \in C} \frac{1}{\mathcal{N}(nB)} \mathcal{H}_2 \left( \frac{\chi}{\mathcal{N}(nB)} \right) = \frac{1}{\mathcal{N}(n)} \sum_{a \in C', \mathcal{N}(a)} \mathcal{H}_2 \left( \frac{\chi}{\mathcal{N}(n)} \right)
\]

which by the orthogonality relations equals

\[
\frac{1}{\mathcal{N}(n)} \sum_{a} \left( \frac{1}{\mathcal{H}_2} \sum_{\chi(H) = 1} \overline{\chi(C')}\chi(a) \right) \frac{1}{\mathcal{N}(a)} \mathcal{H}_2 \left( \frac{\chi}{\mathcal{N}(a)} \right) \frac{1}{\mathcal{N}(n)} \mathcal{H}_2 \left( \frac{\chi}{\mathcal{N}(n)} \right)
\]

\[
= \frac{1}{\mathcal{H}_2\mathcal{N}(n)} \sum_{\chi(H) = 1} \overline{\chi(C')} \sum_{a} \frac{\chi(a)}{\mathcal{N}(a)} \mathcal{H}_2 \left( \frac{\chi}{\mathcal{N}(a)} \right) \frac{1}{\mathcal{N}(n)} \mathcal{H}_2 \left( \frac{\chi}{\mathcal{N}(n)} \right)
\]

so that we can apply the lemma to the inside sums proving the corollary.
It is, perhaps, desirable to justify, at least partially, our choice of "measures" (8), mostly because in the case $K = \mathbb{Q}$ it is customary to use $m$ to measure the size of a congruence class group mod $m$. In keeping with this it seems more appropriate to use $d_K N(m)$ in place of the two "measures" $h_H$, $D_H$, because, if $H$ is a congruence class group mod $m$, then

$$D_H \leq e^{O(n)} d_K N(m) \quad \text{and} \quad h_H \leq e^{O(n)} d_K N(m). \quad (9)$$

For the first of these it suffices to show that if $\chi(H) = 1$ then

$$4^{\omega(m)} - \omega(h_{\chi}) \leq e^{O(n)} N(\chi) \quad \text{which would follow from}$$

$$4^{\omega(m)} - (h_{\chi}) \leq e^{O(n)} \prod_{p/m, p \nmid H_{\chi}} N(p).$$

Letting $v_2, v_3$ denote the number of prime ideals $p$ of $K$ with $Np = 2, Np = 3$ respectively we have $v_2 \leq n, v_3 \leq n$, and

$$4^{\omega(m)} - \omega(h_{\chi}) \leq \left(\frac{4}{2}\right)^{v_2} \left(\frac{4}{3}\right)^{v_3} \prod_{p/m, p \nmid H_{\chi}} N(p),$$

and our contention follows. And, for the second claim we have $h_H \leq (I(m) : P_m)$ which, by p. 127 of Lang [18], is $\leq 2^n h_K \phi(m)$, when we are done by the trivial estimates $\phi(m) \leq N(m)$ and $h_K \leq e^{O(n)} d_K$, the second of which is (21) of I52 (with $a = 1$).

There are three reasons for preferring $h_H, N_H$. The first is that, even when $K = \mathbb{Q}$, the estimates (9) can both be quite weak, as shown by the simple
EXAMPLE: Let $K = \mathbb{Q}$ and let $p_1, p_2, p_3$ be three different primes \( \equiv 1 \mod 4 \), all of roughly the same size. Let $H$ be the class group to the extension field \( L = \mathbb{Q}(\sqrt{p_1 p_2}, \sqrt{p_1 p_3}, \sqrt{p_2 p_3}) \) of $\mathbb{Q}$. Since $p_1, p_2, p_3$ are all of the primes which ramify in $L$, $H$ is a congruence class group mod $\mathfrak{p}^3$ hence "$d_K^N(m)$" = $p_1 p_2 p_3$. On the other hand every $\chi$ so $\chi(H) = 1$ factors through a class group to one of the 3 quadratic subfields of $L$ hence \( D_H = \max(8, 2p_1 p_2, 2p_2 p_3, 2p_1 p_3) \) and $h_H = 4$. If the primes $p_i$ are large the estimate (9) is then quite weak.

The second reason is more formal and is concerned with the notion of what we will call a primitive congruence class group: just as for characters (in I, §1) we say $H_1 \mod m_1$ induces $H_2 \mod m_2$ if $m_1$ divides $m_2$ and if $H_2$ is the preimage of $H_1$ under the canonical homomorphism

\[
\mathcal{I}(m_2)/\mathfrak{p}^{m_2} \rightarrow \mathcal{I}(m_1)/\mathfrak{p}^{m_1}.
\]

Now, in class field theory, one regards $H_1 \mod m_1$ and $H_2 \mod m_2$ as equivalent if there is a congruence class group which induces them both, and denotes the equivalence class by $\overline{H}_1 = \overline{H}_2$. In this context $H_1 \mod m_1$ is called an interpretation (Erklärung) of the equivalence class $\overline{H}_1$ and $m_1$ is the modulus of interpretation. Assigning to a class group $H \mod m$ its conductor

\[
\delta_H = \text{LCM}\{\delta_{\chi} : \chi(H) = 1\}
\]

(10)
it follows, from (81), that \( h_H \) depends only on the equivalence class, that \( \overline{H} \) has an interpretation mod \( \hat{\iota}_H \) (namely the characters satisfying \( \chi(H) = 1 \) are induced by characters mod \( \hat{\iota}_H \)): take the common kernel of these characters mod \( \hat{\iota}_H \) in \( \text{I}(\hat{\iota}_H) \) and that the congruence class groups in the equivalence class \( \overline{H} \) are precisely those congruence class groups induced by the (unique) interpretation of \( \overline{H} \) mod \( \hat{\iota}_H \). We say, then, that a congruence class group \( H \mod m \) is **primitive** if \( m = \hat{\iota}_H \) i.e. \( H \) is already the "smallest" interpretation of \( \overline{H} \).

The importance of the notion of a primitive class group \( H \mod m \) in class field theory is then assured because if \( L/K \) is the class field to \( H \) then \( m \) is also the conductor of the extension; and we get the best bounds for the extension \( L/K \) we can get by the above methods only by considering this primitive congruence class group.

Returning to \( h_H, D_H \) it is clear that \( h_H \) depends only on the equivalence class \( \overline{H} \), but that \( D_H \) does not quite satisfy this condition. For this reason it seems reasonable to complicate matters further and consider instead of \( D_H \) the simpler "measure"

\[
d(H) = \max\{d : \chi(H) = 1\} \tag{11}
\]

where \( d = d_{\chi}N(\hat{\iota}_\chi) \) depends only on the primitive character inducing \( \chi \) (by the definition of \( \hat{\iota}_\chi \)). Then \( d(H) \) also depends only on \( \overline{H} \).

The second reason, then, is that \( h_H, D_H \) (or, better, \( d(H) \)) reflect the properties of the class field to \( H \) more clearly. To justify this we need to show that \( D_H \) and \( d(H) \) are comparable when \( H \) is a
primitive congruence class group, for which purpose we first prove

\[ \frac{1}{h_H} \sum_{\chi(H)=1} \omega(\chi) = \sum_{p/h_H} \left( 1 - \frac{1}{\epsilon_p} \right) \]  

where \( \epsilon_p \) are integers \( \geq 2 \). \hspace{1cm} (12)

But we have

\[ \sum_{\chi(H)=1} \omega(\chi) = \sum_{p/h_H} \#\{ \chi : p/h_H \} = \sum_{p/h_H} \left( h_H - \#\{ \chi : p \nmid h_H \} \right) , \]

and, given \( p/h_H \), the condition \( p \nmid h_H \) means that if \( m \) is the largest divisor of \( h_H \) prime to \( p \) then \( \chi \) (or rather the character mod \( h_H \) it induces) factors through the canonical map \( I(h_H)/P(h_H) \to I(m)/P_m \).

Denoting by \( H_\chi \) the image of \( H \) in \( I(m)/P_m \), it is equivalent to say that \( \chi \) is trivial on the kernel of \( I(h_H)/H \to I(m)/H_\chi \): this kernel has order \( \epsilon_p \geq 2 \), and it is now clear that there are \( h_H/\epsilon_p \) characters \( \chi \) so \( p \nmid h_H \), which proves (12). Incidentally, we remark that if \( L/K \) is class field to \( H \) then \( \epsilon_p \) is the ramification index of (any prime above) \( p \), which explains the notation.

Using (12), we can now prove

If \( H \) is primitive then \( D_H \leq \epsilon \theta(n) d(H)^{1+\epsilon} \) for any \( \epsilon > 0 \). \hspace{1cm} (13)

To see this we put, temporarily, \( F_H = \max\{ N(h_H) : \chi(H) = 1 \} \) and prove the, somewhat stronger, assertion \( 4 \omega(h_H) \leq F_H \epsilon \theta(n) \) for any \( \epsilon > 0 \) (where \( \theta(n) \) means there is a constant \( c_\epsilon \) depending only on \( \epsilon \) so that this quantity has absolute value \( \leq c_\epsilon n \)). It suffices to prove our assertion when \( \epsilon \) is sufficiently small, so we suppose \( \epsilon > 0 \) is small enough so

\[ \log M! > \frac{1}{2} \log M , \text{ for every integer } M > \frac{4}{\epsilon} \]
(which is possible since \( \log M! = M \log M - M + O(\log M) \)). If now \( \omega(h_H) \leq \frac{4}{\varepsilon} + \frac{1}{n} \) then

\[
\frac{\omega(h_H)}{4} \leq \exp\left(4 \log 4 \right) n \leq \frac{\omega(h_H)}{\varepsilon} (n)
\]

and we are done; so we may suppose \( \omega(h_H) > \frac{4}{\varepsilon} + \frac{1}{n} \). Since

\[
\sum_{p \mid h_H} (1 - \frac{1}{p}) \geq \sum_{p \mid h_H} \frac{1}{2} = \frac{1}{2} \omega(h_H), \text{ using (12) allows us to find } \chi \text{ with}
\]

\[
\omega(h_X) \geq \frac{1}{2} \omega(h_H) > \frac{4}{\varepsilon} n. \text{ Fix this } \chi \text{ and put } M = \left[ \frac{\omega(h_X)}{\varepsilon} \right] + 1
\]

\[
> \frac{\omega(h_X)}{n} > 4 \quad \text{(where } [ \ ] \text{ is the greatest integer function): now}
\]

\[
\log N(h_X) \geq \sum_{p \mid h_X} \log Np, \text{ and the } \omega(h_X) \text{ integers } N(p) \geq 2 \text{ coming}
\]

from \( p / h_X \) can repeat no value more than \( n \) times, so there are

\[
\frac{\omega(h_X)}{n} \geq 2 \quad \text{distinct values } \geq 2, \text{ and we have}
\]

\[
\log N(h_X) \geq n \sum_{2 \leq k \leq M} \log k = n \log M! > \frac{n}{2} M \log M >
\]

\[
\frac{1}{2} \omega(h_X) \log \frac{4}{\varepsilon} \geq \frac{\omega(h_H)}{4} \frac{4}{\varepsilon} \log 4 = \frac{\omega(h_H)}{\varepsilon} \log 4.
\]

Thus \( \omega(h_H) < N(h_X)^{\varepsilon} \leq \frac{F}{\varepsilon} \), proving (13).

Actually the second reason is broader than class field theory

because the apparatus of L-series also reduces everything to primitive

characters (by 3) of I§1): thus studying the L-series \( L(s, \chi) \) for \( \chi \)

with \( \chi(H) = 1 \) reduces, in some sense, to the case where \( H \) is

primitive, as will be seen in the paragraph preceding theorem 4.
Finally the third reason has to do with our eventual application of these results to prime ideals: for if \( H_1, H_2 \) are equivalent in the above sense then there are only finitely many prime ideals whose placement in the cosets \( \text{mod } H_1 \) or \( \text{mod } H_2 \) is changed by changing the interpretation modulus. Because we are approaching the distribution of prime ideals via the distribution of integral ideals (by Selberg's sieve method) where the interpretation modulus makes a difference (at least analytically) it makes sense to minimize this effect. For example, in the proof of (14) in IV52 we will see how prime ideals can be handled more effectively than integral ideals. The applications however will not come until chapter V.

§3. Application of Selberg's method

We need, for our main application, an analogue of the corollary to lemma 8 (with \( n = 0 \)) which sums the prime ideals instead of all integral ideals. In keeping with Selberg's method we define, for \( z \geq 1 \),

\[
S_z = \text{the set of integral ideals which have no prime ideal factor } p \text{ with } Np \leq z
\]

(1)

and shall find it convenient to express our result in the terms of the quantity

\[
V(z) = \sum_{Na \leq z} \frac{1}{Na}
\]

(2)

to which we devoted chapter II. We can state
LEMMA 9. Let $C$ be a coset of a congruence class group $H \mod m$. Then for $y > 0$ and $z \geq 1$ we have

$$\sum_{\alpha \in C \cap S_z} \frac{1}{N(a)} H_{2n}(\frac{Y}{Na}) \leq \frac{\kappa(K)}{h_H V(z)} + O(e^{\frac{0(n)}{a n (D_H^{1/2} A^{n/2} z^2)}^{1+1/a} A y^{-1}}),$$

for any $a \geq 2$ and $A \gg 1$.

PROOF: This is a consequence of Selberg's sieve with the main estimating being done by the corollary to lemma 9, which we take in the following form (mainly for convenience):

$$\sum_{\alpha \in C \cap n/a} \frac{1}{Na} H_{2n}(\frac{Y}{Na}) = \frac{1}{N(n)} \frac{\phi(m)}{N(m)} \frac{\kappa(K)}{h_H} +$$

$$O(e^{\frac{0(n)}{a n (D_H^{1/2} A^{n/2})^{1+1/a} A y^{-1}}}),$$

which follows from (7) of §2, or also from the statement of the corollary, since

$$\left(\log D_H A^n\right)^n \leq \left(\frac{1}{e^a} \frac{\log D_H A^n}{2n}\right)^n = a^n (D_H A^n)^{2a}$$

(recall that the log terms came originally by choosing $a$ to minimize this expression; this was done mainly to show that this term is quite small).

Let now $\lambda \in e$ be real numbers, one for each integral ideal $\mathfrak{c}$, and satisfying
i) \( \lambda_e = 0 \) unless \( e \) is squarefree, prime to \( m \), with 
\[ Ne \leq z \]

ii) \( \lambda_0 = 1 \) where \( 0 \) is the unit ideal

iii) \( |\lambda_e| \leq 1 \) for all \( e \),

but otherwise arbitrary. If \( a \in \mathbb{Z}_z \) and \( e \) divides \( a \), then either \( e = 0 \) or \( \lambda_e = 0 \); from this follows

\[
\sum_{a \in S^c} \frac{1}{Na} H_{2n} \left( \frac{y}{Na} \right) \leq \sum_{a \in \mathbb{C}} \frac{1}{Na} H_{2n} \left( \frac{y}{Na} \right) \left( \sum_{e/a} \lambda_e \right)^2
\]

(4)

because \( H_{2n} \) is non-negative, by (2) and (10) of §1. Noting that all but finitely many \( \lambda_e \) are 0, and writing \([e_1, e_2], (e_1, e_2)\) for the LCM, respectively GCD, of \( e_1 \) and \( e_2 \), we write the right side of (4) as

\[
\sum_{e_1, e_2} \lambda_{e_1} \lambda_{e_2} \sum_{a \in \mathbb{C}} \frac{1}{N(a)} H_{2n} \left( \frac{y}{Na} \right) = \sum_{[e_1, e_2]: \text{divides } a} [e_1, e_2]
\]

by (3), since \([e_1, e_2]\) is indeed prime to \( m \) by condition i),

whenever \( \lambda_{e_1} \lambda_{e_2} \neq 0 \). By \( e_1 e_2 = [e_1, e_2](e_1, e_2) \) and i), iii), this becomes

\[
\frac{\phi(m)}{N(m)} \frac{\kappa(K)}{h} \sum_{e_1, e_2} \frac{\lambda_{e_1} \lambda_{e_2}}{Ne_1 Ne_2} N(e_1, e_2) + O(e^0(n) a^{1 + 1/a} Ay^{-1}) \sum_{Ne_1 \leq z} \sum_{Ne_2 \leq z}
\]
\begin{align*}
&= \frac{\phi(m)}{N(m)} \frac{\kappa(K)}{h_H} \sum_{e_1, e_2} \lambda e_1 \lambda e_2 \sum_{a e_2} \phi(a) + O\left(e^{O(n)} a^{n \left(\frac{1}{2} \frac{n}{H} A^2\right) \frac{1+1/a}{a} A_y^{-1} \left(\sum_{N \in \mathbb{Z}^{+}} \right)^2} \right) \\
&= \frac{\kappa(K)}{h_H} \frac{\phi(m)}{N(m)} \sum_{a} \phi(a) \sum_{e_1, e_2} \lambda e_1 \lambda e_2 + O\left(e^{O(n)} a^{n \left(\frac{1}{2} \frac{n}{H} A^2\right) \frac{1+1/a}{a} A_y^{-1} \left(\sum_{N \in \mathbb{Z}^{+}} \right)^2} \right) \\
&\quad \text{divisible by } a
\end{align*}

by (20) of (5.2). Going back to (4) we now have

\begin{align*}
\sum_{a \in \mathbb{Z}^+, \eta \in \mathbb{C}} \frac{1}{Na} H_2(n \frac{y}{Na}) \leq \frac{\kappa(K)}{h_H} \frac{\phi(m)}{N(m)} \sum_{a} \phi(a) \left(\sum_{a e_2} \lambda e_2 \right)^2 + O\left(e^{O(n)} a^{3n \left(\frac{1}{2} \frac{n}{H} A^2\right) \frac{1+1/a}{a} A_y^{-1}} \right)
\end{align*}

(5)

where the error term is as claimed, and where the \( \lambda e \) are any real numbers subject to \( i), ii), iii) \). Clearly, it suffices to show that there exist \( \lambda e \), compatible with these conditions, with

\begin{align*}
\frac{\phi(m)}{N(m)} \sum_{a} \phi(a) \left(\sum_{a e_2} \lambda e_2 \right)^2 \leq \frac{1}{V(z)} ,
\end{align*}

(6)

which is the heart of Selberg's method: to construct the \( \lambda e \)'s we find real numbers \( \beta_a \) satisfying
a) \( \beta_a = 0 \) unless \( a \) is squarefree, prime to \( m \), with \( Na \leq z \)

b) \( \sum \mu(a)\beta_a = 1 \)

c) \( \sum \phi(a)\beta_a^2 = 1/\nu_m(z) \)

where \( \mu = \mu_K \) is the Möbius function of \( K \) and

\[
\nu_m(z) = \sum_{a \text{ prime to } m} \frac{\mu^2(m)}{\phi(m)}
\]  \hspace{1cm} (7)

Then defining

\[
\lambda_e = Ne \sum B \mu(B)\beta_{Be}
\]  \hspace{1cm} (8)

we verify that these \( \lambda_e \) do everything we need; this is just a change of variables, for we then have

\[
\sum_{a \neq e} \frac{\lambda_e}{Ne} = \beta_a
\]  \hspace{1cm} (9)

since

\[
\sum_{a \neq e} \sum B \mu(B)\beta_{Be} = \sum_{e_1 B} \sum \mu(B)\beta_{B(ae_1)} = \sum_D B e_1 \frac{\sum \mu(B)\beta_{Be_1}}{Be_1 = D}
\]

\[
= \sum_D \left( \sum_{B \neq D} \beta_{aD} \right) = \beta_a .
\]

The reason for making the change of variables is that the conditions a), b), c) can be considered as a minimum problem (before \( \nu_m(z) \) has even been defined): namely, minimize

\[
\sum \phi(a)\beta_a^2 \text{ subject to } a) \text{ and } b). \]

Using Lagrange multipliers, for example, this leads to the choice
\[
\beta_a = \begin{cases} \frac{\mu(a)}{\phi(a)\nu_m(z)} & \text{if } Na \leq z, \ a \text{ is prime to } m \\ 0 & \text{otherwise} \end{cases}
\]

for which \( a, b, c \) are obvious. Defining \( \lambda_e \) by (8) we must verify that i), ii), iii) hold: if \( e \) is not squarefree, or not prime to \( m \), or \( N_e \) is not \( \leq z \), then the same holds of \( ae \) for any integral ideal \( a \), hence \( \beta_{ae} = 0 \) for all \( a \), and so \( \lambda_e = 0 \) by (8), verifying i). The condition (ii) follows immediately from (b), so it remains to verify iii): we may suppose \( e \) is prime to \( m \), and squarefree, when

\[
\frac{\lambda_e}{N_e} \sum_B \mu(B) \beta_{Be} = \sum_{NBe \leq z/N_e} \mu(B) \frac{\mu(Be)}{\nu_m(z)\phi(Be)}
\]

\[
= \frac{1}{\nu_m(z)} \sum_{NBe \leq z/N_e} \frac{\mu(B)\mu(Be)}{\phi(Be)} \quad \text{since } e \text{ was prime}
\]

\( B \) prime to \( m \) and \( e \) to \( m \), and since \( \mu(Be) = 0 \) unless \( B \) is prime to \( e \). Since \( \mu, \phi \) are multiplicative and \( \mu(e) = \pm 1 \) we have

\[
\nu_m(z) | \lambda_e | = \frac{N_e}{\phi(e)} \sum_{NBe \leq z/N_e} \frac{\mu^2(B)}{\phi(B)}
\]

\( B \) prime to \( Ne \)

and need the identity

\[
\frac{N(e)}{\phi(e)} = \sum_{a | e} \frac{\mu^2(a)}{\phi(a)}
\]  

(10)
which holds because each side equals \( \prod_{p \mid e} (1 + \frac{1}{\phi(p)}) \). Then by (10)

we get

\[
V_m(z) |\lambda_e| = \sum_{a \mid e} \sum_{NB \leq z / Ne} \frac{\mu^2(aB)}{\phi(aB)} \leq V_m(z)
\]

since, clearly, only some of the terms in the sum \( V_m(z) \) are represented. It now remains only to verify that (6) is satisfied

i.e. by (9) and c), we need only check that \( \frac{\phi(m)}{N(m)} \frac{1}{V_m(z)} \leq \frac{1}{V(z)} \).

By using the identity (10) again, we have

\[
\frac{N(m)}{\phi(m)} V_m(z) = \sum_{B \mid m} \sum_{NBSz / Ne} \frac{\mu^2(Ba)}{\phi(Ba)} \geq \sum_{Ne \leq z} \frac{\mu^2(e)}{\phi(e)}
\]

a prime to \( m \)

since any squarefree \( \epsilon \) with \( Ne \leq z \) can be expressed as \( Ba \).

But, denoting by \( \sigma(B) \) the largest squarefree divisor of an ideal \( B \), we have, for square free ideals,

\[
\frac{1}{\phi(e)} = \frac{1}{Ne} \prod_{p \mid e} \frac{1}{1 - 1/Np} = \frac{1}{Ne} \prod_{p \mid e} \left( 1 + \frac{1}{Np} + \frac{1}{Np^2} + \ldots \right) = \sum_{\sigma B = e} \frac{1}{NB} ;
\]

hence

\[
\frac{N(m)}{\phi(m)} V_m(z) \geq \sum_{Ne \leq z} \frac{1}{\phi(e)} = \sum_{Ne \leq z} 1 = \sum_{\sigma B = e} \frac{1}{NB} ;
\]

\[
\sum_{Na \leq z} \frac{1}{Na} = V(z) ,
\]
since if $Na \leq z$ then $\sigma a$ has norm $\leq z$. With this the proof of lemma 9 is complete.

§4. Conclusion

Finally we combine the formal mean value theorem of §1 with the distribution statements (on the average) of §2 (respectively §3) for ideals in generalized arithmetic progressions (respectively prime ideals). To avoid repeating hypothesis we fix a complex-valued function $b(a)$ defined on the integral ideals of $K$ and satisfying

$$\sum_{\alpha} |b(\alpha)| < \infty \quad (1)$$

and a congruence class group $\mathbb{H} \mod m$. Our first result generalizes theorem 2 of Gallagher [10] to algebraic number fields.

**Theorem 3.** If $b, H$ are as above and $T \gg 1$ then

$$\sum_{\chi(\mathbb{H})=1} \int_{-T}^{T} \left| \sum_{\alpha} b(\alpha) \chi(\alpha)N(\alpha)^{-it} \right|^2 dt \ll$$

$$\sum_{\alpha} |b(\alpha)|^2 (\kappa(\mathbb{K})N(\alpha) + e^{0(n)} \frac{n}{2} h H \frac{1}{H} \frac{1}{n} \frac{1}{n+1} (\log \frac{D^n T}{H} ))^n$$

**Proof:** We apply lemma 7' with $\ell = 2n$ and $A = (2n)^2 T$ to the term associated to $\chi$ obtaining
\[
\int_{-T}^{T} \left| \sum_{a} b(a) \chi(a) N(a) e^{-it} \right|^2 dt \ll \int_{0}^{\infty} \left| \sum_{a} b(a) \chi(a) H_{2n} \left( \frac{\nu}{Na} \right) \right|^2 dy
\]

(2)

for each \( \chi \), formally putting \( b_m = \sum_{Na=m} b(a) \chi(a) \) in lemma 7', and noting the hypotheses there follow from (1). Observe that the parameter \( A \) chosen to define the \( H_{2n} \) functions is the same for each \( \chi \), and that

\[
\left| \sum_{\chi(H)=1} a \left| \sum_{a} b(a) \chi(a) H_{2n} \left( \frac{\nu}{Na} \right) \right|^2 = h_H \sum_{\mathbb{C} \in \Pi(m)/H} \left| \sum_{a \in \mathbb{C}} b(a) H_{2n} \left( \frac{\nu}{Na} \right) \right|^2 \right.
\]

(3)

since on squaring the left side we get

\[
\sum_{\chi} \sum_{a_1, a_2} b(a_1)b(a_2) H_{2n} \left( \frac{\nu}{Na_1} \right) H_{2n} \left( \frac{\nu}{Na_2} \right) \chi(a_1) \chi(a_2)
\]

where there are no convergence problems because all these sums are finite by (2) and (10) of §1. By the orthogonality relations this becomes

\[
\sum_{a_1 \equiv a_2 \bmod I(m)/H} b(a_1)b(a_2) H_{2n} \left( \frac{\nu}{Ha_1} \right) H_{2n} \left( \frac{\nu}{Ha_2} \right) h_H
\]

\[
= h_H \sum_{\mathbb{C} \in \Pi(m)/H} \sum_{a_1 \in \mathbb{C}} \sum_{a_2 \in \mathbb{C}} b(a_1)b(a_2) H_{2n} \left( \frac{\nu}{Na_1} \right) H_{2n} \left( \frac{\nu}{Na_2} \right)
\]

which is the right side of (3).
Now summing (2) on $x$, interchanging $\sum_x$ with $\int_0^\infty$ and applying (3) leaves us with

$$\sum_x \int_1^T \left| \int b(a) x(a) N(a) - \frac{t}{2} \right|^2 dt \lesssim \sum_{h \in I(m)/H} \int_0^\infty \int C(a) H \left( \frac{y}{Na} \right)^2 dy$$

(4)

on also interchanging $\sum_C$ with $\int_C$. Applying the Cauchy-Schwartz inequality shows that for each $C$

$$\left| \sum_{a \in C} b(a) H_{2n} \left( \frac{y}{Na} \right) \right|^2 \leq \sum_{a \in C} |b(a)|^2 N(a) H_{2n} \left( \frac{y}{Na} \right) \sum_{a \in C} \frac{1}{Na} H_{2n} \left( \frac{y}{Na} \right)$$

(5)

noting once again that the $H_{2n}$ are non-negative. We no longer concern ourselves with convergence issues because all terms of sums and integrals will be non-negative from now on, so we can consider divergence as convergence to infinity: in this case our conclusions have no interest but are nevertheless correct (actually in the applications all the sums will be finite). The point, of course, is that the last sum in (5) is handled by the corollary to lemma 8 which gives

$$\sum_{a \in C} \frac{1}{Na} H_{2n} \left( \frac{y}{Na} \right) = \frac{\phi(m)}{N(m)} \frac{\kappa(K)}{h_h} + O(M) y^{-1}$$

(6)

where $M = e^{O(n)} \frac{1}{n} \frac{1}{2} \frac{1}{2} \frac{1}{2} \log \frac{D_H^N T}{M}$ abbreviates the error term (with $A = (2n)^{1/2} T$) of the corollary, independent of $y$. Writing $S$ for the expression on the left hand side of the statement of the theorem
and combining (4), (5), (6) shows that

\[ S \ll \sum_{a \in I(m) / H} \int_0^\infty \sum_{a \in C} |b(a)|^2 N(a) H_2 n \left( \frac{\phi(m)}{N(m)} \kappa(K) \right) + O(HM)y^{-1} \frac{dy}{y} \]  

(7)

\[ = \sum_{a \in I(m)} |b(a)|^2 \frac{\phi(m)}{N(m)} \kappa(K) N(a) \int_0^\infty H_2 n \left( \frac{\phi(m)}{N(m)} \kappa(K) \right) N(a) \int_0^\infty \frac{dy}{y} \frac{dy}{y} \]

so we now need to know that

\[ \int_0^\infty H_2 n \left( \frac{\phi(m)}{N(m)} \kappa(K) \right) N(a) \int_0^\infty \frac{dy}{y} \frac{dy}{y} = m^{-\alpha} h_{2n}(\alpha) \]  

(8)

which we see by making the substitution \( y = \exp u \), \( \frac{dy}{y} = u \) so our integral becomes

\[ \int_{-\infty}^\infty H_2 n \left( \exp u \right) m^{-\alpha} e^{-au} du = m^{-\alpha} \int_{-\infty}^\infty h_{2n}(u) e^{-au} du = m^{-\alpha} h_{2n}(\alpha) \]

by (2) and (4) of §1.

Noting that \( h_{2n}(0) = 1 \) and that \( |h_{2n}(1)| \leq \theta(n) \) by (13) of §1 and \( h_{2n}(s) = h_{2n}(s) \) we find that (7) and (8) combine to give

\[ S \ll \sum_{a \in I(m)} |b(a)|^2 \left[ \frac{\phi(m)}{N(m)} \kappa(K) N(a) + O(\theta(n)MH) \right] \]

which is still sharper than what is claimed.
Before continuing we note that although the appearance of \( n^n \) was successfully avoided in §2, §3 it has again surfaced in theorem 3 via lemma 7'; in fact, an alternate treatment would avoid \( n^n \) in lemma 7' but then it seems again to be forced on us in the treatment of §2, §3.

We come finally to the main result of this chapter which improves theorem 3 for sums with prime ideal arguments. More precisely, we prove

**Theorem 3'.** If \( b, H \) are as in (1), if \( T \gg 1 \) and if \( b(a) = 0 \) whenever \( a \) has a prime ideal factor \( p \) with \( Np \leq z \) then

\[
\sum_{x(H) = 1}^{T} \left| \sum_{a} b(a) x(a) N(a)^{-it} \right|^2 dt \ll 
\]

\[
\ll \sum_{a} |b(a)|^2 \left( \frac{k(\chi)}{V(z)} N(a) \right) + e^{\theta(n)} \frac{n}{a^2} 3n \left( \frac{1}{\sqrt{T}} \right)^2 \sum_{a_H = 1}^{T} \frac{1}{a_H T}
\]

for any \( \alpha \geq 2 \), where

\[
V(z) = \sum_{Na \leq z} \frac{1}{Na}
\]

**Proof:** We follow the proof of theorem 3, replacing the use of the corollary to lemma 8 by the use of lemma 9. Thus we put again

\[
A = (2n)^{1/2} T \quad \text{and} \quad \ell = 2n \quad \text{and repeat the argument of theorem 3 up to (4) without change. At this stage we make the only change in the proof: namely recalling the definition of } S_z \text{ in §3 we observe that the sum on the right side of (4) equals}
\]
since \( b(a) \) vanishes outside \( S_z \) by hypothesis. This allows us to replace (5) by
\[
\left| \sum_{a \in \mathbb{C} \cap S_z} b(a) H_{2n}(\frac{y}{Na}) \right|^2 \leq \sum_{a \in \mathbb{C}} |b(a)|^2 N(a) H_{2n}(\frac{y}{Na}) \sum_{a \in \mathbb{C} \cap S_z} \frac{1}{Na} H_{2n}(\frac{y}{Na}) \tag{9}
\]
hence we can use lemma 9 to replace (6) by
\[
\sum_{a \in \mathbb{C} \cap S_z} \frac{1}{Na} H_{2n}(\frac{y}{Na}) \leq \frac{k(K)}{h_H y(z)} + O(M) y^{-1} \tag{10}
\]
this time with
\[
M = e^{O(n)} \frac{n}{a} 3n (D_H 1/2 a/2 2^{1/2} a_T)
\]
on putting \( A = (2n)^{1/2} T \) in the error term of lemma 9 and using \( a \geq 2 \).
Writing again \( S \) for the expression on the left side of the statement of theorem 3' and combining (4), (9), (10) as before, we obtain
\[
S \ll \sum_{a \in I(m)} |b(a)|^2 \left[ \frac{k(K)}{V(z)} N(a) \int_0^\infty H_{2n}(\frac{y}{Na}) \frac{dy}{y} + O(h_H M) Na \int_0^\infty H_{2n}(\frac{y}{Na}) y^{-1} \frac{dy}{y} \right] \tag{11}
\]
so that applying (8) and the estimates \( h_{2n}(0) = 1 \) and \( |h_{2n}(1)| \leq e^{O(n)} \) as before we finally have.
\[ S \ll \sum_{a \in I(m)} |b(a)| \left[ \frac{\kappa(n)}{V(z)} N(a) + O(e^{0(n)} n^{-H_M}) \right] \]

which is again more than sufficient. It may appear that we have lost more in theorem 3 by suppressing the \( \phi(m)/N(m) \) term (which can be quite small) than we have in theorem 3', however this is illusory since a similar loss is involved in the transition from \( V_m(z) \) to \( V(z) \) in the proof of lemma 9.

To use theorem 3' it is necessary to recall that by the first part of the corollary to lemma 5 (with \( E = K \)) we have

\[ \frac{\kappa(n)}{V(z)} \ll \frac{1}{\log z} \quad \text{provided} \quad z \geq c_n^{3n_d K} \quad (12) \]

without which theorem 3' is useless. Note that even if we had been able to avoid \( n^n \) in this chapter the use of (12) reintroduces it.
CHAPTER IV
THE ZERO-DENSITY THEOREM

Finally we are ready to prove the main result, which shows that the number of zeros of an L-function near the line $\sigma = 1$ is small. The method is that used by Fogels, Gallagher and Bombieri for this purpose, each author improving the method over the previous (the last two for $K = \mathbb{Q}$). At the critical stages in §2, where they are needed, we apply the result of the previous chapters, which were designed with this purpose in mind. We maintain our usual emphasis on $K$-uniformity, however this will require very little work at this stage because our tools have been adequately prepared.

§1. Detecting non-exceptional zeros

Using the results of §2,§3, of chapter III and purely analytical arguments we show that, in the presence of a non-exceptional zero, a certain arithmetic sum must be "large". The argument used was discovered by Fogels [7] in exactly our present situation and is based on an idea of Turán, which we quote as lemma 10. Fogels, however, uses a different arithmetic sum as his indicator of the presence of a zero (interestingly, based on the "averaging functions" of chapter III although for very different values of the arguments: this suggests that the arguments of this § can be combined with those of chapter III directly) and also a different technique for using the "largeness" of the indicator (than what will be used in §2) and consequently his method will not work near the pole at $s = 1, \chi = 1$; also he does not attempt to control the dependence of his constants on the degree
of the field (in particular they grow with the degree as follows from the Introduction). Based on Fogel's work, Gallagher [10] introduces the simpler "indicator" used here and also exploits its' size via an integration (as we shall in §2); the simpler "indicator" allows us to avoid some of the complications which arise in Fogel's method (although his is more general), and Gallagher's integration technique allows us, eventually to make Turán's lemma yield information near the pole at $s = 1, \chi = 1$. Finally, Bombieri [2] uses Gallagher's formulation and is able to show that the zero density theorem becomes sharper when the exceptional zero exists. Both Gallagher and Bombieri restrict themselves to $K = \mathbb{Q}$ so do not have to pay attention to the pole at $s = 1, \chi = 1$ (because $\mathbb{Q}$-uniformity = absolute uniformity).

For the general case we must deal with this pole, especially when the exceptional zero exists (if it doesn't, Fogels finds a different, simpler argument to deal with the pole), which we do (via the $\Phi_\chi$ term in lemma 11) by Turán's method. This is the only novel feature in §1.

As in §3, let $Q \geq 1, T \geq 1$ satisfy $Q^T \gg 1$ and put $L = \log QT^n$. Define

$$e_1 = \begin{cases} 1 & \text{if the exceptional zero } (\rho_1, \chi_1) \text{ for } Q, T \text{ exists} \\ 0 & \text{otherwise} \end{cases}$$

$$L_1(\psi) = \begin{cases} L(\psi) \setminus \{\rho_1\} & \text{if } e_1 = 1 \text{ and } \psi = \chi_1 \\ L(\psi) & \text{if not} \end{cases}$$

for any primitive character $\psi$. Note that $L_1(\psi)$ is the set (with multiplicities) of non-exceptional zeros of $L(s, \psi)$ whenever $d_\psi \leq Q$.
(although we define \( \mathcal{L}_1(\psi) \) also when \( d_\psi > Q \)). We regard \( Q, T \) as fixed in this chapter so our notation does not reflect the dependence of \( e_1 \) and \( \mathcal{L}_1(\psi) \). Also, we should point out that the condition \[ \frac{1}{Q^T} \gg 1 \] is used to imply that \( Q^T \) is large enough so that all the results of \( \S 3 \) hold, large enough that

\[ \rho_1 > \frac{3}{4}, \text{ if } \rho_1 \text{ exists} \quad (2) \]

and, finally, large enough that lemma 11 (when it appears) is not vacuous.

Let now \( \chi \) be a primitive character and \( \tau \) a real number so that

\[ d_\chi \leq Q, \quad |\tau| \leq T. \quad (3) \]

Then, when \( \mathcal{L}^{-1} \ll \tau \ll 1 \) (a condition we state more precisely later), we are concerned with detecting non-exceptional zeros of \( L(s, \chi) \) which satisfy

\[ |\rho - (1+i\tau)| \leq \tau, \quad \rho \in \mathcal{L}_1(\chi). \quad (4) \]

The method is based on the study of

\[ F(s, \chi) = \frac{L'}{L}(s, \chi) + e^{\frac{L'}{1L}(s+1-\rho_1, \chi \chi_1)} \quad (5) \]

when \( s \) is near \( 1+i\tau \), the \( e_1 \) part eventually leading to sharper results when \( e_1 = 1 \) (the idea of Bombieri). We begin by showing that in the disc \( |s - (1+i\tau)| \leq \frac{1}{4} \) we have
where $G(s, \chi)$ is analytic and $\ll \mathcal{L}$ in the disc.

This is basically lemma 4 but there are some details to attend to: first $\chi \chi_1$ is not necessarily primitive, so letting $\chi \chi_1$ be induced by the primitive character $\psi$ we will apply lemma 4 to $L(s, \psi)$ and then return to $L(s, \chi \chi_1)$ by

$$\frac{-L'}{L}(s+1-\rho_1, \chi \chi_1) = \frac{-L'}{L}(s+1-\rho_1, \psi) + \sum_{p/\mathcal{H}_1 \mathcal{H}_1, \mathcal{H}_1 \mathcal{H}_1} \frac{\psi(p) \log Np}{N(p)^{s+1-\rho_1-\psi(p)}}$$

(as in the proof of (3) in I§3) where we need only observe that $|s-1+it| \leq \frac{1}{4}$ implies $\operatorname{Re}(s+1-\rho_1) \geq \operatorname{Re}(s) > 3/4$ allows us to estimate the sum (very weakly) as

$$\ll \sum_{p/\mathcal{H}_1 \mathcal{H}_1} \frac{\log Np}{N(p)^{3/4-1}} \ll \sum_{p/\mathcal{H}_1 \mathcal{H}_1} \log N(p) \leq \log N(\mathcal{H}_1 \mathcal{H}_1) \leq 2 \log Q \ll \mathcal{L}.$$ 

Also we have, in (6), written $\mathcal{L}_1(\chi \chi_1)$ in place of $\mathcal{L}_1(\psi)$ and will, shortly, write $\delta_{\chi \chi_1}$ for $\delta_{\psi}$.

Now, from $|s-(1+i\tau)| \leq 1/4$ follows

$$|s-(1+i\tau)| \leq \frac{1}{2} \quad \text{and} \quad |(s+1-\rho_1) - (1+i\tau)| \leq |1-\rho_1| + |s-(1+i\tau)| \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$
so lemma 4 applies to both terms in (5) after the transition from \( XX_1 \) to \( \psi \) of the last paragraph. Thus

\[
F(s, \chi) = -\frac{\delta_X}{s-1} + \sum_{\rho \in \mathcal{L}(\chi)} \frac{1}{s-\rho} \left( \frac{e_1 \delta_{XX_1}}{(s+1-\rho_1)-1} + \sum_{\rho \in \mathcal{L}(XX_1)} \frac{e_1}{s+1-\rho_1-\rho} + G(s, \chi) \right)
\]

holds in \(|s-1-i\tau| < 1/4\) with \( G(s, \chi) \) equal to the transition sum of the last paragraph plus the error terms for \( \chi, \psi \) of lemma 4 (note that \( d_\psi \leq Q^2 \) not \( Q \) so "\( \mathcal{L} \)" from lemma 4 will be \( \mathcal{L} \)).

It thus remains only to change from sums over \( \mathcal{L}(\chi), \mathcal{L}(XX_1) \) to those over \( \mathcal{L}_1(\chi), \mathcal{L}_1(XX_1) \): but there is an exceptional zero only if \( e_1 = 1 \), and then it appears in \( \mathcal{L}(\chi) \) only when \( \chi = \chi_1 \) (equivalently \( \delta_{XX_1} = 1 \)) while in \( \mathcal{L}(XX_1) \) only when \( \psi = \chi_1 \) (equivalently \( \delta_\chi = 1 \)).

So we have

\[
F(s, \chi) = -\frac{\delta_X}{s-1} + \frac{e_1 \delta_{XX_1}}{s-\rho_1} - \frac{e_1 \delta_{XX_1}}{s-\rho_1} + \frac{e_1 \delta_\chi}{s+1-\rho_1-\rho_1} + \text{the claimed sums over } \mathcal{L}_1(\chi), \mathcal{L}_1(XX_1) + G(s, \chi)
\]

which proves (6)–the fact that the \( 1/s-\rho_1 \) terms cancel is very important.

Before continuing with (6) we quote the relevant result of Turán as

**Lemma 10.** If \( z_1, \ldots, z_m \) are complex numbers and \( K \geq m \) then there exists \( k \) with \( K \leq k \leq 2K \) and

\[
|z_1^k + \ldots + z_m^k| \geq (2c_5 |z_1|)^k
\]

where \( c_5 \) is an absolute constant with \( 0 < c_5 < \frac{1}{2} \).
This is proved in Sos, Turán [26], where it is shown that
\[ c_5 = \frac{1}{8} e^{-1-4/e} > 1/96 \] is admissible, as a special case of a more flexible result.

To apply lemma 10 we put
\[ \xi = 1 + r + i\tau, \quad \xi_1 = \xi + (1 - \rho_1) \quad \text{with} \quad r > 0 \] \[ (7) \]
then apply the operator \( \frac{(-1)^k}{k!} \frac{d^k}{ds^k} \) to (6) and put \( s = \xi \) so obtaining
\[ \frac{(-1)^k}{k!} \frac{d^k}{ds^k} F(\xi, x) = \delta \left( \frac{e_1}{(\xi - \rho_1)^{k+1}} - \frac{1}{(\xi - 1)^{k+1}} \right) + \sum_{\rho \in \mathcal{L}_1} \frac{1}{(\xi - \rho)^{k+1}} \frac{1}{|\rho - 1 - i\tau| < 1} \]
\[ + \sum_{\rho \in \mathcal{L}_1} \frac{1}{(\xi - \rho)^{k+1}} \frac{1}{|\rho - 1 - i\tau| < 1} \]
\[ \text{for} \quad 0 < r < 1/8 \]
where only the error term needs proof: since \( r < 1/8 \), \( G(s, x) \) is analytic in a region containing the disc \( |s - \xi| \leq \frac{1}{8} \) (since there \( |s - 1 - i\tau| \leq |s - \xi| + r < \frac{1}{8} + \frac{1}{8} = \frac{1}{4} \) hence, by the Cauchy integral formulas,
\[ \frac{(-1)^k}{k!} \frac{d^k}{ds^k} G(\xi, x) = \frac{(-1)^k}{2\pi i} \int_{|s - \xi| = \frac{1}{8}} \frac{G(s, x)}{(s - \xi)^{k+1}} ds \]
when (8) follows immediately from the estimate (6) for \( G \).

It is evident how lemma 10 applies to (8), making precise our remark preceding lemma 4. This will ultimately lead us to
**Lemma 11.** Let $Q \geq 1$, $T \geq 1$ satisfy $Q^nT \gg 1$ and put $L = \log QT^n$.

There exist effective constants $c_{11}, c_{12}, c_{13}, c_{14}, c_{15}, c_{16} > 0$ independent of our field $K$ so that for any $y, r$ satisfying

$$y \geq e^{c_{11}L}, \quad c_{15}^{y-1} \leq r \leq c_{16} (< 1/8)$$

the following holds: if $\chi$ is a primitive character (on $K$) and $\tau$ is a real number so that

$$d_{\chi} \leq Q, \quad |\tau| \leq T$$

and if there exists a non-exceptional zero $\rho$ (for $T, Q$) with $|\rho - (1 + i\tau)| \leq r, \rho \in L_1(\chi)$ then

$$\phi(\tau, r)^2 + r^4\log y \sum_{y^\epsilon} \left( \prod_{y < \rho \leq u} \frac{\gamma(p)\chi(p)}{1+\tau} \right)^2 \frac{du}{u} \gg y^{-c_{14}r}$$

where we have put

$$\gamma(p) = \left( 1 + e_1 \frac{\chi_1(p)}{1-\rho_1} \right) \log N(p)$$

for prime ideals $\rho$ of $K$, where $e_1$ is defined by (1), and where

$$\phi(\tau, r) = \begin{cases} 1 & \text{if } e_1 = 0, \chi = 1, |\tau| < 2c_5^{-1}r \\ y^{c_{13}r} (1-\rho_1)^{1/2} r^{-1/2} & \text{if } e_1 = 1, \chi = 1, |\tau| < 2c_5^{-1}r \\ 0 & \text{in all other cases} \end{cases}$$
PROOF: First a word about the constants: they will appear with increasing suffix, so those not mentioned yet (namely $c_6, c_7, c_8, c_9, c_{10}$) will occur in the course of the proof and then the constants in the statement of the lemma will be constructed in terms of them. Except for $c_{15}, c_{16}$, each of those constants will be fixed as soon as it appears; for $c_{15}, c_{16}$ we suppose at the outset that $c_{15}^{r-1} \leq r \leq c_{16}$ and then by increasing $c_{15}$ (decreasing $c_{16}$) finitely many times to satisfy the requirements of the moment they will be determined only at the end.

We begin by fixing $x, \tau$ as in the statement of the lemma, and showing that if $C > 1$ is chosen appropriately then (8) becomes, for $k \geq 1$,

$$
\frac{(-1)^k}{k!} \frac{d^k}{ds^k} F(\xi, x) = \delta \left( \frac{e_1}{(\xi_1 - \rho_1)^{k+1}} - \frac{1}{(\xi - 1)^{k+1}} \right) + \sum_{\rho \in \mathcal{L}_1} (\xi - \rho)^{-k-1} (9)
$$

$$
+ \sum_{\rho \in \mathcal{L}_2} (\xi_1 - \rho)^{-k-1} + O((Cr)^{-k}) ,
$$

provided $c_2C^{-1}r^{-1} < r < \frac{1}{8}c^{-1}$ and a result which will allow $m$ in lemma 10 to be chosen rather small.

Proving this amounts to diminishing the range of summation in (8) by using lemma 3: considering $\mathcal{L}_1(x)$ we choose $j_o$ so

$$
2^{j_o}Cr \leq 1 < 2^{j_o+1}Cr
$$

when

$$
\sum_{\rho \in \mathcal{L}_1(x),Cr \leq |\rho - 1i\tau| < 1} (\xi - \rho)^{-k-1} \leq \sum_{0 \leq j < j_o} \sum_{2^{j}Cr \leq |\rho - 1i\tau| < 2^{j+1}Cr} (\xi - \rho)^{-k-1}
$$
\[ + \sum_{\mathbf{1} \leq \mathbf{c} \leq |\rho - 1 - i\tau| \cdot 1} |\xi - \rho|^{-k-1} \lesssim \sum_{0 \leq j < j_0} (2^j \mathbf{c} \mathbf{r})^{-k-1} \sum_{2^j \mathbf{c} \leq |\rho - 1 - i\tau| < 2^{j+1} \mathbf{c} \mathbf{r}} 1 \]

\[ + (2^j \mathbf{c} \mathbf{r})^{-k-1} \sum_{2^j \mathbf{c} \leq |\rho - 1 - i\tau| < 1} 1 \]

since \[ |\xi - \rho| = [(r + 1 - \Re \rho)^2 + (\tau - \Im \rho)^2] \geq |1 + i\tau - \rho| \] by \( r \geq 0 \).

Now, by lemma 3, the number of zeros of \( L(s, x) \) satisfying \( |\rho - 1 - i\tau| < R \) is \( \ll R \mathcal{L} \) provided that \( 1 \geq R > c_2 \mathcal{L}^{-1} \); thus provided \( \mathbf{c} \mathbf{r} > c_2 \mathcal{L}^{-1} \), as stipulated in (9), we will have

\[ \ll (\mathbf{c} \mathbf{r})^{-k} \sum_{\rho \in \mathcal{L}^{-1}(x), \mathbf{c} \mathbf{r} \leq |\rho - 1 - i\tau| < 1} |\xi - \rho|^{-k-1} \lesssim \sum_{0 \leq j < j_0} (2^j \mathbf{c} \mathbf{r})^{-k-1} \cdot 2^j \mathcal{L} + (2^j \mathbf{c} \mathbf{r})^{-k-1} \mathcal{L} \]

\[ \ll (\mathbf{c} \mathbf{r})^{-k} \sum_{j=0}^{j_0} 2^{-j \mathbf{c} \mathbf{r}} \ll (\mathbf{c} \mathbf{r})^{-k} \mathcal{L} \] where the \( j_0 \) term fits in the sum

because \( 1 < 2 \sum_{j=0}^{j_0+1} \mathbf{c} \mathbf{r} \). The argument for \( \mathcal{L}(\chi \chi_1) \) is similar although we should point out that if \( \chi \chi_1 \) is induced by the primitive character \( \psi \) (as before) then \( d_\psi \leq Q^2 \) (not \( Q \)) so that \( \mathcal{L} \) is replaced by \( \mathcal{L} \) and the lower bound for \( |\xi - \rho| \) used before is replaced by \( |\xi_1 - \rho_1| = [(r + 1 - \rho_1 + 1 - \Re \rho)^2 + (\tau - \Im \rho)^2]^{1/2} \geq |1 + i\tau - \rho| \), which holds already by \( r + 1 - \rho_1 \geq 0 \). Finally \( r < \frac{1}{8} \mathcal{L}^{-1} \) guarantees \( 8^k < (\mathbf{c} \mathbf{r})^{-k} \) so that the error term of (8) can be absorbed in \( \mathcal{O}((\mathbf{c} \mathbf{r})^{-k} \mathcal{L}) \), and (9) is proved.

Now, by lemma 3 again the number of terms in the sums over \( \mathcal{L}(x), \mathcal{L}(\chi \chi_1) \) is
where \( c_6 \) incorporates the constants for \( \chi \) and \( \chi \chi_1 \) which are left implicit in lemma 3 (which applies because \( Cr^{L} > c_2 \)). Suppose that a non-exceptional zero \( \rho_o \) of \( L(s, \chi) \) satisfying \( |\rho_o - (1+it)| \leq r \) actually does exist. Then by (9), (10) and lemma 10:

if \( K \geq c_6Cr^{L} \) then there is a \( k \) with \( K \leq k \leq 2K \) and

\[
\left| \frac{1}{k!} \frac{d^k}{ds^k} F(\xi, \chi) \right| + \delta \left| \frac{1}{(\xi-1)^{k+1}} - \frac{e_1}{(\xi-\rho_o)^{k+1}} \right| \geq \left( \frac{2c_5}{|\xi-\rho_o|} \right)^{k+1} \tag{11}
\]

Continuing from (11) we observe \( |\xi-\rho_o| \leq r + |1+it-\rho_o| \leq 2r \) so that the right hand side of (11) is

\[
\geq c_5^{k+1} r^{-k-1} (1-((c_5C)^{-k} r^{L}))
\]

Fix \( C > c_5^{-1} \): then \( c_5C > 1 \) so by (11) we have

\[
(c_5C)^{-k} r^{L} \leq (c_5C)^{-k} r^{L} \leq (c_5C)^{-k} r^{L} = \theta r^{L} r^{L}
\]

where \( \theta = (c_5C)^{-k} < 1 \). Thus by insisting \( r^{L} \) is large (ie. that \( c_{15} \) is large enough), and putting

\[
c_7 = c_6 C
\]

we deduce from (11) that

\[
\leq c_6(Cr^{L})
\]
if \( K \geq c_7 x^{\gamma} \) then there is a \( k \) with \( K \leq k \leq 2K \)

and

\[
\frac{1}{k!} \frac{d^k}{d s^k} F(\xi,\chi) + \delta \chi \left( \frac{1}{\xi^{-1})^{k+1} + \frac{e_1}{(\xi^{-1}_r)^{k+1}} \right) \geq \frac{3}{4} c_5^{k+1} r^{-k-1}.
\]

We next deal with the \( \delta \chi \) term and first observe that if

\[ |\tau| \geq 2 c_5^{-1} r \]

then

\[ |\xi -\xi| \geq |\tau| \geq 2 c_5^{-1} r \]

and \( |\xi -\rho_1| \geq |\tau| \geq 2 c_5^{-1} r \)

(since \( \xi -\xi \) and \( \xi -\rho_1 \) have imaginary part \( \tau \) hence

\[
\left| \frac{1}{\xi^{-1})^{k+1} + \frac{e_1}{(\xi^{-1}_r)^{k+1}} \right| \leq 2 \frac{c_5^{-1} r}{2 r} = 2^{-k} c_5^{-k+1} r^{-k-1} < \frac{1}{12} c_5^{-k+1} r^{-k-1}
\]

provided \( k \geq 4 \) (which is certain if \( c_{15} \) is large enough since by (11): \( k \geq K \geq c_7 x^{\gamma} \geq c_7 c_{15} \)). Thus we have proved most of

if \( K \geq c_7 x^{\gamma} \) then there is a \( k \) with \( K \leq k \leq 2K \)

and

\[
\frac{1}{k!} \frac{d^k}{d s^k} F(\xi,\chi) + \delta \chi \phi_k(\tau, r) \geq \frac{2}{3} c_5^{k+1} r^{-k-1}
\]

where for suitable \( c_8 > 1 \),

\[
\phi_k(\tau, r) = \begin{cases} 1 & \text{if } e_1 = 0, |\tau| < 2 c_5^{-1} r \\ c_8 (1 - \rho_1)^{1/2} r^{-1/2} & \text{if } e_1 = 1, |\tau| < 2 c_5^{-1} r \\ 0 & \text{if } |\tau| \geq 2 c_5^{-1} r \end{cases}
\]
which is a provisional \( \phi_X \) (which will follow by removing the dependence on \( k \)). Clearly it remains to show

\[
\left| \frac{1}{(\xi - 1)^{k+1}} - \frac{e_1}{(\xi - \rho_1)^{k+1}} \right| \leq \phi_k(\tau, r) r^{-k-1} \text{ when } |\tau| < 2c_5^{-1} r ;
\]

suppose first that \( e_1 = 0 \) : then

\[
\left| \frac{1}{(\xi - 1)^{k+1}} \right| = \left| \frac{1}{(r + i\tau)^{k+1}} \right| \leq r^{-k-1} \text{ as required. So it remains to consider } e_1 = 1; \text{ then we put } e_1 = 1 - \rho_1 \text{ temporarily and recall that } \rho_1 \text{ exceptional means (by lemma 2)}
\]

\[
e_1 \leq 2c_1 \tau^{-1} \leq \frac{1}{2} c_1 5^{\tau} \leq \frac{r}{2} \text{ provided } c_1 5 \geq 4c_1 ,
\]

as we may suppose. By the calculus of residues and (7):

\[
\frac{1}{(\xi - 1)^{k+1}} - \frac{1}{(\xi - \rho_1)^{k+1}} = \frac{1}{(r + i\tau)^{k+1}} - \frac{1}{(r + i\tau + 2\varepsilon_1)^{k+1}}
\]

\[
= \frac{1}{2\pi i} \int_{s-(r + i\tau + \varepsilon_1)}^{s-(r + i\tau) + 1/2} \frac{1}{s^{k+1}} \left( \frac{1}{s-(r + i\tau)} - \frac{1}{s-(r + i\tau + 2\varepsilon_1)} \right) ds
\]

because the disc \( |s - r - i\tau - \rho_1| \leq \varepsilon_1^{1/2} r^{1/2} \) contains \( s = r + i\tau \) and \( s = r + i\tau + 2\varepsilon_1 \) but not \( s = 0 \) (all by \( \varepsilon_1 < r \)). Now since

\[
e_1^{1/2} r^{1/2} = |s - r - \varepsilon_1 - i\tau| \leq |s - r - 2\varepsilon_1 - i\tau| + \varepsilon_1 \text{ and}
\]

\[
e_1^{1/2} r^{1/2} = |s - r - \varepsilon_1 - i\tau| \leq |s - r - i\tau| + \varepsilon_1 \text{ and}
\]

\[
r \leq |r + \varepsilon_1 + i\tau| \leq |s| + |r + \varepsilon_1 + i\tau - s| = |s| + \varepsilon_1^{1/2} r^{1/2}
\]
on the circle \(|s-(r+i\tau+\epsilon_1)| = \epsilon_1^{1/2} r^{1/2}\) we have

\[
\left| \frac{1}{(\xi-1)^{k+1}} - \frac{1}{(\xi_1 - \rho_1)^{k+1}} \right| \leq \int_{|s-r-\epsilon_1-\epsilon| = \epsilon_1^{1/2} r^{1/2}} \frac{ds}{|s|^{k+1} |s-r-i\tau| |s-r-i\tau-2\epsilon_1|} \cdot \epsilon_1^{1/2} \frac{2\pi \epsilon_1^{1/2} r^{1/2}}{|r-r_1^{1/2}-r_1^{1/2}|^{k+1}} \frac{\epsilon_1^{1/2} r^{1/2}-\epsilon_1^{1/2}}{(r^{1/2} - \epsilon_1^{1/2})^{k+3}} = \frac{2\epsilon_1^{1/2} r^{-k/2}}{(r^{1/2} - \epsilon_1^{1/2})^{k+3}}
\]

because \(r^{1/2} - \epsilon_1^{1/2} \geq r^{1/2}(1-\frac{1}{\sqrt{2}})\) and because \(k\) can be supposed large by increasing \(c_{15}\) if necessary (as we have observed earlier). Thus (13) is proved.

Returning to the \(F\) term in (13) where \(F\) is defined in (5)) we use the Dirichlet series representation (16) of \(I_{s2}\) to write

\[-F(s,\chi) = \sum_{\alpha} \frac{\chi(\alpha)\chi(\alpha)}{\alpha} N(\alpha)^s \text{ for } s > 1\]

where
\[ \gamma(a) = \left( 1 + \frac{e^{\frac{1}{2a}}}{N(a)^{1-\rho_1}} \right) \Lambda(a) \quad \text{has} \quad |\gamma(a)| \leq 2\Lambda(a) \quad (15) \]

generalizes the function in the statement of the lemma to all (integral) ideals. Applying the operator \( \frac{(-1)^k}{k!} \frac{d^k}{ds^k} \) to the series for \(-F\) and putting \( s = \xi \) (with \( \text{Re}(\xi) = 1 + r > 1 \)) then gives

\[ \frac{(-1)^{k+1}}{k!} \frac{d^k}{ds^k} F(\xi, \chi) = \frac{1}{k!} \sum_{a \in N(a)^{1+r+i\tau}} \frac{\gamma(a) \chi(a)}{(\log Na)^k} \quad (16) \]

which, on defining

\[ j_k(u) = \frac{e^{-u}}{k!} \quad , \quad u > 0 \quad (16) \]

becomes

\[ \frac{(-1)^{k+1}}{k!} \frac{d^k}{ds^k} F(\xi, \chi) = \sum_{a \in N(a)^{1+r+i\tau}} \frac{\gamma(a) \chi(a)}{(\log Na)^k} j_k(r \log Na) \quad (17) \]

Our next task is to show that in (17) only the terms with a prime and \( r \log Na \) of order of magnitude \( k \) (where \( j_k(u) \) has a maximum) can contribute to (13). To do this we find \( c_9, c_{10} \) so that

\[ j_k(u) \leq \begin{cases} \left( \frac{e}{2} \right)^k, & u \leq c_9^k \\ \left( \frac{e}{2} \right)^k e^{-u/2}, & u \geq c_{10}^k \end{cases} \quad (18) \]

But \( \frac{k^k}{k!} \leq e^k \) so \( j_k(u) \leq \left( \frac{eu}{k} \right)^k e^{-u} \leq \left( \frac{eu}{k} \right)^k \leq \left( \frac{e}{2} \right)^k \) for \( u \leq \frac{c_5}{2e} \)

and it suffices to take any \( c_9 \leq \frac{c_5}{2e} \); also \( j_k(u)e^{u/2} = \frac{u^k e^{-u/2}}{k!} \)
has a maximum at $u = 2k$ so choosing $c_{10} > 2$ with $c_{10}^{e^{1-\frac{1}{2}c_{10}}} \leq \frac{c_5}{2}$ we find that

$$j_k(u)e^{u/2} \leq \frac{(c_{10}^k)^{k} e^{-c_{10}^k/2}}{k!} \leq \left(\frac{e}{c_9}\right)^k (c_{10}^k)^{k} e^{-c_{10}^k/2} = (c_{10}^{e^{1-\frac{1}{2}c_{10}}})^k$$

$$\leq \left(\frac{c_5}{2}\right)^k$$

as (18) claims.

We also want to free ourselves from the dependence of $k$ so we put

$$K = c_9^{-1} r \log y$$

(19)

and the condition (13) on $K$ is met whenever $y \geq e^{c_7 c_9 y^L}$ which is ensured by taking (and fixing)

$$c_{11} \geq c_7 c_9.$$

Then by (13),

$$c_9^{-1} r \log y \leq k \leq 2c_9^{-1} r \log y$$

(20)

so (18) implies

$$j_k(r \log N_\alpha) \leq \begin{cases} 
\left(\frac{c_5}{2}\right)^k, & N_\alpha \leq y \\
\left(\frac{c_5}{2}\right)^{k - \frac{1}{2}r}, & N_\alpha \geq y^{c_{12}} 
\end{cases}$$

(21)

where we have put

$$c_{12} = 2c_9^{-1} c_{10}.$$
To verify (21) we note that:

from $Na \leq y$ follows $r \log Na \leq r \log y \leq c_9 k$ by (20), and

from $Na \geq y$ follows $r \log Na \geq c_{12} r \log y \geq \frac{1}{2} c_{12} c_9 k$ by (20)

and $\frac{1}{2} c_{12} c_9 k = c_{10} k$. In either case we can apply (18) with $u = r \log Na$.

We are now prepared to eliminate the low order terms from (17), for which purpose the principal tool is (18), (19) of I§2; the hypotheses of these results is that $r$ is small enough and that $y$ is large enough, which we can ensure by insisting that $c_{16} \leq \frac{1}{2} c_0$ (so $2r$ is small enough) and that $\mathcal{L} \gg 1$ (so $y \geq e^{c_{11}}$, with fixed $c_{11}$, implies $y$ is large enough).

By (15), (21) and by (19) of I§2 we get

$$\left| \sum_{Na \leq y} \frac{\gamma(a)(a)}{N(a)^{1+1/\tau}} j_k(r \log Na) \right| \leq 2 \left( \frac{c_5}{2} \right)^k \sum_{Na \leq y} \frac{\Lambda(a)}{N(a)} \ll \left( \frac{c_5}{2} \right)^k \log d_k y$$

$$\ll \left( \frac{c_5}{2} \right)^k \log y \ll \left( \frac{c_5}{2} \right)^k \frac{k}{r} \text{ because } \log d_k \leq \log Q \leq \mathcal{L} \ll \log y$$

(and $c_{11}$ has been fixed) and because $\log y \ll \frac{k}{r}$ by (20). Now $\frac{1}{2} c_5$ is certainly $< c_5$ so if $k$ is large enough then

$$\left| \sum_{Na \leq y} \frac{\gamma(a)(a)}{N(a)^{1+1/\tau}} j_k(r \log Na) \right| < \frac{1}{12} c_5 \frac{k+1}{r-1} :$$

(22)

as before we can ensure $k \gg r \log y \gg r \mathcal{L}$ is large by raising $c_{15}$ again (if necessary).

Similarly, by (15), (21) and by (18) of I§2 we have
| \[ \frac{\gamma(a)\chi(a)}{N(a)^{1+\frac{1}{2}}} j_k(r \log Na) \leq \left( \frac{c_5}{2} \right)^k \sum_{a \neq \pm 1} \frac{\Lambda(a)}{N(a)^r} \leq \left( \frac{c_5}{2} \right)^k \left( \frac{1}{r + \ell} \right) \]

| \[ \leq (k+1) \left( \frac{c_5}{2} \right)^k \frac{1}{r-1} \], since \( \ell \ll \log y \ll \frac{k}{r} \) as in the proof of (22); this is also the same kind of estimate as in the proof of (22) so if \( k \) is large enough (by raising \( c_{15} \) again) then

| \[ \left( \frac{c_5}{2} \right)^k \frac{1}{r} \cdot \left( k+1 \right) \left( \frac{c_5}{2} \right)^k \frac{1}{r-1} \]. \quad (23) \]

Finally we eliminate the non-primes from (17) by using

| \[ j_k(u) \leq \sum_{\ell=0}^{\infty} j_k(u) = 1 \] \quad (24) \]

which is immediate from the definition (16); from (24) we deduce

| \[ \frac{j_k(r \log Na)}{Na^{1/2}} = \left( \frac{2r}{k!} \frac{\frac{1}{2} \log Na}{\Lambda(a)} \right)^k e^{\frac{1}{2} \log Na} \]

and then have

| \[ \left( \frac{c_5}{2} \right)^k \sum_{a \neq \pm 1} \frac{\Lambda(a)}{N(a)^r} \leq \left( \frac{c_5}{2} \right)^k \left( \frac{1}{r + \ell} \right) \]

| \[ \leq (2r)^k \sum_{a=\pm \rho, m\geq 2} \frac{\Lambda(a)}{N(a)^{1+2r}} \]

| \[ = (2r)^k \sum_{\rho} \log N\rho \sum_{m=2}^{\infty} \frac{1}{m^{1+r}} = (2r)^k \sum_{\rho} \frac{\Lambda\rho}{N\rho^{1+2r}} \left( 1 - \frac{1}{N\rho^{1/2+r}} \right)^{-1} \]

| \[ \leq (2r)^k \sum_{a=\pm \rho, m\geq 2} \frac{\Lambda(a)}{N(a)^{1+2r}} \leq (2r)^k \left( \frac{1}{r} + \ell \right) \quad \text{by (18) of I\S2 again. But} \]
\[ r \leq c_{16} \quad \text{and} \quad \frac{1}{r} + \mathcal{L} \ll (k+1)r^{-1} \quad \text{(as in proof of (23)) so our sum is} \]
\[ \ll (k+1)(2c_{16})^k r^{-1} \ll (k+1)\left(\frac{c_{16}}{2}\right)^k r^{-1} \]
on choosing \( c_{16} \leq \frac{1}{4}c_5 \). This has been dealt with before so, if \( c_{15} \) is raised appropriately, we get

\[
\sum_{y \leq Na \leq y} \frac{y(a) \chi(a)}{N(a)^{1+\tau}} j_k(r \log Na) < \frac{1}{12}c_5 \frac{k+1}{r^{-1}} \quad \text{(25)}
\]

once again. Note that \( c_{16} \) can now be fixed: the only demands on \( c_{16} \) have been \( c_{16} < \frac{1}{8}c^{-1} \) in (8) and \( c_{16} \leq \frac{1}{2}c_0 \), \( c_{16} \leq \frac{1}{4}c_5 \) just above.

Now multiplying (13) by \( r^k \), combining it with (17), and taking account of (22), (23), (25) brings us to

\[
\sum_{y \leq Np \leq y} \frac{y(p) \chi(p)}{N(p)^{1+\tau}} j_k(r \log Np) + r^{-1} \phi_{\chi}^2(r, r) > \frac{5}{12}c_5 \frac{k+1}{r^{-1}} \quad \text{(26)}
\]

for some \( k \) satisfying \( c^{-1}_9 r \log y \leq k \leq 2c^{-1}_9 r \log y \). Here we have used the transition from \( K \) to \( y \) expressed by (19) and (20) and must still verify \( \delta_{\chi}^k \phi_{\chi}^* \leq \phi_{\chi}^k \): but this is immediate since

\[ c_k = e^{k \log c_8} \leq e^{(2c^{-1}_9 \log c_8)r \log y} \leq y^{c_{13}r} \]

where we have used (20) and \( c_8 > 1 \) (see (14)) and taken \( c_{13} \geq 2c^{-1}_9 \log c_8 \) (> 0).
Defining

\[ S(u) = \sum_{y < N \leq u} \frac{\gamma(p) \chi(p)}{N(p)^{1+\tau}}, \quad y \leq u \leq y^{c_{12}} \]

the sum of the statement of the lemma, and summing by parts shows

\[ \sum_{y < N \leq u} \frac{\gamma(p) \chi(p)}{N(p)^{1+\tau}} j_k(r \log Np) = S(y^{c_{12}}) j_k(r \log y^{c_{12}}) \]

\[ - \int_y^{y^{c_{12}}} S(u) j'_k(r \log u) \frac{r}{u} du \]

since \( S(y) = 0 \). By (21) again, we find

\[ |S(y^{c_{12}}) j_k(r \log y^{c_{12}})| \ll \left( \frac{c_5}{2} \right)^k \frac{1}{2^{c_{12}r}} \sum_{y < N \leq y^{c_{12}}} \frac{\Lambda(p)}{Np} \]

\[ \ll \left( \frac{c_5}{2} \right)^k \frac{1}{r \log y} \ll \left( \frac{c_5}{2} \right)^k \frac{r^{-1}}{y^{2c_{12}r}} \gg r \log y \]

and (19) of §2, because \( \log d_K \ll y \ll \log y \). As before if \( k \) is large enough (hence if \( c_{15} \) is again increased) this shows that

\[ |S(y^{c_{12}}) j_k(r \log y^{c_{12}})| < \frac{1}{12c_{15}^k} r^{-1} \]

which together with (26), (27) shows that

\[ r^{-1} \phi(r, r) + \int_y^{y^{c_{12}}} |S(u)| \frac{du}{u} > \frac{1}{3c_{15}^k} r^{-1} \]  \( (28) \)
for some $k$ satisfying (20); here we have used

$$|j_k'(u)| = |j_{k-1}(u) - j_k(u)| \leq j_{k-1}(u) + j_k(u) \leq 1$$

as in (24). Observe that the left side of (28) is already free of $k$; we free the right side by putting $c_{14} = 4c_9^{-1} \log \frac{1}{c_5}$ which is $> 0$ since $c_5^{-1} > 2$ (see lemma 10) and then noting that

$$c_5^{-k} = e^{-(\log \frac{1}{c_5})k}, \quad -c_5^{-1} \log \frac{1}{c_5} \log y \leq c_5^{-k} \leq e^{(-2c_9^{-1} \log \frac{1}{c_5})r \log y} \leq \frac{1}{2} c_{14}r$$

by using (20).

Then we square both sides of (28) and use

$$2(x^2 + y^2) \geq (x + y)^2$$

and

$$\left( \int_y^y S(u) \frac{du}{u} \right)^2 \leq \int_y^y |S(u)|^2 \frac{du}{u} \leq \log y \int_y^y S(u)^2 \frac{du}{u}$$

by the Cauchy-Schwartz inequality: putting these steps together expresses (28) as

$$r^{-2} \phi_x(r) + r^2 \log y \int_y^y |S(u)|^2 \frac{du}{u} \gg y^{-c_{14}r} r^{-2}$$

which is the statement of the lemma (finally!) so we can let our interminable increasing of $c_{15}$ come to a stop.
§2. The main theorem

Combining results of all the previous sections (except Theorem 1 which has so far been used only in a non-essential way - see remarks after theorem 2) will enable us to prove our main result. We first fix our field $K$ as in chapter I and fix some notation.

Given $\alpha, Q, T$ so that $Q \geq 1, T \gg 1, Q^\frac{1}{T} \gg 1$ and $\frac{1}{2} < \alpha < 1$ we put

$$L = \log QT^n$$

$$e_1 = \begin{cases} 1 & \text{if } (\rho_1, \chi_1) \text{ exists (for } Q, T) \\ 0 & \text{if not} \end{cases} \quad (1)$$

$$A_1 = \begin{cases} (1-\rho_1)L, & \text{if } e_1 = 1 \\ 2c_1, & \text{if } e_1 = 0 \end{cases}$$

and define for characters $\chi$ satisfying $d_\chi \leq Q : N_1(\alpha, T, Q; \chi) =$ the number of non-exceptional zeros $\rho$ of $L(s, \chi)$ which satisfy

$$|\Im \rho| \leq T, \quad \Re \rho \geq \alpha \quad (2)$$

Moreover if $H$ is a congruence class group with $d(H) \leq Q$ put

$$N_1(\alpha, T, Q; H) = \sum_{\chi(H)=1} N_1(\alpha, T, Q; \chi) \quad (3)$$

where we recall (from III§2) that

$$d(H) = \max\{d_\chi : \chi(H) = 1\}$$

$$d_\chi = d_\chi N(\rho_\chi) \quad (4)$$
It is appropriate to make a few remarks about these normalizations: first, by lemma 2, we have \( \Lambda_1 \leq 2c_1 \) in both cases so the "more exceptional" \( \rho_1 \) is, the smaller is \( \Lambda_1 \); second, we note that if \( \chi, \psi \) are characters which are induced by the same primitive character then \( d_\chi = d_\psi \) and \( N_1(\alpha, T, Q; \chi) = N_1(\alpha, T, Q; \psi) \) (since \( \delta_\chi = \delta_\psi \) and \( L(s, \chi), L(s, \psi) \) have different zeros only on \( \sigma = 0 \), by (3) of I§1; we are not concerned with zeros on \( \sigma = 0 \) by \( \alpha > 1/2 \)). The restriction \( d_\chi \leq Q \) on definition (2) is needed so that the notion of exceptional zero is well-defined (see I§4) and the restriction \( d(H) \leq Q \) on definition (3) is then as weak as possible. By these remarks we see that \( N_1(\alpha, T, Q; H) \) depends only on the equivalence class \( \overline{H} \) of \( H \) (in the sense of III§2) as does the invariant \( d(H) \): this amplifies the "second reason" of III§2.

Our goal is

**THEOREM 4** Let \( \theta, \lambda > 0 \) be given, let \( Q, T, \alpha \) be given as in (1) except that in addition to \( Q^T n > 1 \) we require \( Q^T n \leq n \) and \( c_{17} \leq \alpha < 1 \). If now \( H \) is a congruence class group with \( d(H) \leq Q \) and \( h_H \leq (QT^n)^\lambda \) then we have

\[
N_1(\alpha, T, Q; H) \ll \Lambda_1(Q^n)^{c_{18}(1-\alpha)}
\]

where the implied constant for \( \ll \) depends only on \( \lambda, \theta \).
PROOF: The parameters \( \lambda, \theta \) are included because there are no clear indications on what values to put there. We write \( QT^n = e^{\mathbb{L}} \) throughout and put

\[
c_{17} = 1 - \frac{c_{16}}{\sqrt{2}}
\]

noting that \( c_{16} > \frac{1}{8} \) (see lemma 11) makes \( c_{17} \) (much) larger than \( 1/2 \).

We begin by applying lemma 11 to a single (primitive) character \( \chi \) with \( d \chi \leq Q \) putting

\[
r = \sqrt{2}(1-\alpha)
\]

and letting \( y \geq e^{c_{11}\mathbb{L}} \) be arbitrary, yet. The choice of \( c_{17} \) above shows lemma 11 holds provided

\[
\alpha \leq 1 - \frac{c_{15}}{\sqrt{2}}
\]

a condition we accept until the end of the proof. Let \( C \) denote a constant to be determined: we will prove that various inequalities hold for all \( y \geq e^{C\mathbb{L}} \) provided \( C \geq c_{11} \) (and may later have to increase \( C \)).

We denote by \( \rho \) a general non-exceptional zero of \( L(s, \chi) \) satisfying

\[
\text{Re } \rho \geq \alpha , \quad |\text{Im } \rho| \leq T
\]

and observe that if \( \tau \) is a real number with \( |	au| \leq T \) and if \( \rho \) is such a zero then

\[
|\text{Im } \rho - \tau| \leq \frac{T}{\sqrt{2}} \implies |\rho - (1+i\tau)| \leq r ,
\]
since if $|\text{Im } \rho - \tau| \leq r/\sqrt{2}$ then

$$|\rho - (1+it)|^2 = (\text{Re } \rho - 1)^2 + (\text{Im } \rho - \tau)^2 \leq (\alpha - 1)^2 + \frac{r^2}{2} + \frac{r^2}{2} = r^2$$

by (5).

For each $\rho$ as above define a function $\psi_{\rho}$ by

$$\psi_{\rho}(\tau) = \begin{cases} 1 & \text{if } |\tau-\text{Im } \rho| \leq \frac{r}{\sqrt{2}} \text{ for } |\tau| \leq T \\ 0 & \text{if not} \end{cases} \quad (9)$$

Moreover by $\sum_{\rho}$ we denote the sum over such zeros where each one occurs to the multiplicity that it occurs in $L(s,\chi)$. Then we claim

$$\sum_{\rho} \psi_{\rho}(\tau) \ll r \mathcal{L} \quad \text{for } |\tau| \leq T. \quad (10)$$

But $\sum_{\rho} \psi_{\rho}(\tau) = \text{the number of non-exceptional zeros of } L(s,\chi) \text{ (counting multiplicities) satisfying}$

$$|\text{Im } \rho - \tau| \leq \frac{r}{\sqrt{2}} , \text{ Re } \rho \geq \alpha ,$$

which by (8) is $\ll$ the number of zeros (counting multiplicities) of $L(s,\chi)$ satisfying $|\rho - (1+it)| \leq r$. By lemma 3 this number is indeed $\ll r \mathcal{L}$ provided $r \geq c_2^{-\mathcal{L}-1}$ (which holds by (5) and (6) since $c_{15} > c_2$ by (9) of §1).

On the other hand, (8) shows that if

$$\psi_{\rho}(\tau) = 1 \text{ then } |\rho - (1+it)| \leq r$$

hence lemma 11 applies and shows
for $|\tau| \leq T$, for if $\psi_\rho(\tau) = 0$ then this is trivial.

But again if $\rho$ is a zero as in (7) then

$$\int_{-T}^{T} \psi_\rho(\tau) d\tau \geq \frac{r}{\sqrt{2}}$$

since the interval $[-T,T] \cap [\text{Im } \rho - \frac{r}{\sqrt{2}}, \text{Im } \rho + \frac{r}{\sqrt{2}}]$ has length $\geq \frac{r}{\sqrt{2}}$

because of $|\text{Im } \rho| \leq T$. Thus, integrating (11) over $\tau$ and summing over the zeros $\rho$ (with multiplicity) shows that

$$y^{-c_14r} \frac{r}{\sqrt{2}} N_1(\alpha, T, Q; \chi) \leq y^{-c_14r} \sum_{\rho} \int_{-T}^{T} \psi_\rho(\tau) d\tau
$$

$$\ll \sum_{\rho} \int_{-T}^{T} \psi_\rho(\tau) \left( \phi_\chi(\tau, r)^2 + r^4 \log y \int_{y}^{y^{c_{12}}} \left| \sum_{y < N\rho \leq u} \frac{\gamma(p) x(p)}{N(p)^{1 + i\tau}} \right|^2 \frac{du}{u} \right) d\tau
$$

$$= \int_{-T}^{T} \left( \sum_{\rho} \psi_\rho(\tau) \right) \left( \phi_\chi(\tau, r)^2 + r^4 \log y \int_{y}^{y^{c_{12}}} \left| \sum_{y < N\rho \leq u} \frac{\gamma(p) x(p)}{N(p)^{1 + i\tau}} \right|^2 \frac{du}{u} \right) d\tau
$$

$$\ll rL \int_{-T}^{T} \left( \phi_\chi(\tau, r)^2 + r^4 \log y \int_{y}^{y^{c_{12}}} \left| \sum_{y < N\rho \leq u} \frac{\gamma(p) x(p)}{N(p)^{1 + i\tau}} \right|^2 \frac{du}{u} \right) d\tau
$$

by
\[ N_1(a, T, Q; x) \ll \delta \chi \Lambda y \left( C_{14} + 2 C_{13} \right) T \]

where it remains only to show

\[ \mathcal{L} \int_{-T}^{T} \phi_\chi (\tau, r)^2 d\tau \ll \delta \chi \Lambda y \]

since one \( r \) on each side of the original inequality cancels. By the definition of \( \phi_\chi \) in lemma 11 we see that (13) is obvious unless \( \chi = 1 \), which we assume, and then we must consider the cases \( e_1 = 0 \) and \( e_1 = 1 \) separately. If \( e_1 = 0 \) then

\[ \mathcal{L} \int_{-T}^{T} \phi_\chi (\tau, r)^2 d\tau = \mathcal{L} \int_{-2c_5^{-1}r}^{2c_5^{-1}r} 1 d\tau \ll \mathcal{L} \ll r \log y \ll 2c_{13}^r \]

\[ e_1 = y \]

as required since \( \Lambda_1 \gg 1 \) in this case. On the other hand if \( e_1 = 1 \) then

\[ \mathcal{L} \int_{-T}^{T} \phi_\chi (\tau, r)^2 d\tau = \mathcal{L} \int_{-2c_5^{-1}r}^{2c_5^{-1}r} \frac{1}{y (1 - \rho_1)^{r-1}} d\tau \ll (1 - \rho_1) \mathcal{L} \]

\[ \ll 2c_{13}^r \]

again as required since \( \Lambda_1 = (1 - \rho_1) \mathcal{L} \) in this case.
We now consider $H$ subject to $d(H) \leq Q$, as in the statement of the theorem, and observe that, by the remarks preceding the statement of the theorem, we may suppose $H$ is a primitive congruence class group (in the sense of III§2). We want to apply (11) to all the characters $\chi$ with $\chi(H) = 1$: these $\chi$ are, however, not necessarily primitive (contrary to the hypothesis of lemma 11). On the other hand, the sum in (12) is over primes; suppose $\chi(H) = 1$, so $\chi$ is defined mod $\delta_H$ and the only difference between the expression (12) for $\chi$ and for the primitive character $\chi^*$ inducing it (on putting $\delta_\chi = \delta_{\chi^*}$) is in the values $\chi(p)$ and $\chi^*(p)$ for those primes $p$ dividing $\delta_H$. However if $p/\delta_H$ then $p/\delta_\psi$ for some $\psi$ so $\psi(H) = 1$ (by (10) of III§2), since $H$ is primitive) hence $N(p) \leq N(\delta_\psi) \leq Q$. If, then, we insist that $C \geq 1$ then, by $y \geq e^{4c_1} \geq Q$, these primes $p$ are not in the sums (21) anyway: thus (12) applies to all $\chi$ with $\chi(H) = 1$, provided $C \geq 1$.

Summing (12) over $\chi$ with $\chi(H) = 1$ then yields

$$N_1(\alpha, T, Q; H) \ll \Delta_1 \frac{(c_{14} + 2c_{13})r}{y}$$

$$(14)$$

$$N_2(\alpha, T, Q; H) \ll \Delta_2 \frac{(c_{14} + 2c_{13})r}{y} \sum_{y < N \leq T} \left\{ \sum_{1 \leq \nu \leq \delta_H} \frac{\chi(p)}{\nu} \right\}^2 \frac{du}{u}$$

since $\delta_\chi = 0$ for all but one $\chi$ and where we have used that

$r \log y \ll e^{2c_{13}} \log y = \frac{1}{y} c_{13}^r$ implies $\frac{c_{14} r}{y} \leq \log y =$


To estimate the remaining, complicated looking expression we use the results of chapter II and III: first, by theorem 3' with \( a = 3 \), \( z = y^{1/3} \) and

\[
\begin{align*}
\theta \left( \frac{y}{N(p)} \right) & \text{ if } a = p \text{ prime, } y < Np < u \\
0 & \text{ otherwise}
\end{align*}
\]

we find that

\[
\sum_{x(H)=1}^{T} \left| \sum_{y<Np<u} \frac{\theta(y)}{N(p)^2} \right|^2 dt
\]

\[
\ll \sum_{y<Np<u} \frac{|\theta(y)|^2}{N(p)^2} \left( \frac{\lambda(K)}{V(y^{-1/3})} N(p) + e^{O(n) n/2} \left( \frac{n}{D_H T} 2/3 \log y \right)^{4/3} \right)
\]

\[
\ll \frac{1}{\log y} \sum_{y<Np<u} \frac{|\theta(y)|^2}{N(p)} \left( 1 + e^{O(n) n/2 D_H 2/3 T} \frac{2}{3} n^{1+H_H y} 8/9 \log y \right)
\]

by (12) of III§4, provided \( y^{1/3} \geq c_3 n^{3n d_k} \).

Now

\[
\begin{align*}
c_3 n^{3n d_k} & \leq e^{O(n) (Q^n)^9/9^{9/3}} \leq e^{O(L) + 9L + 3L} \leq e^{O(L)}
\end{align*}
\]

since \( n \ll L \) by \( Q^n T \gg 1 \): thus if \( C \) is large enough then \( y \geq e^{C L} \) will indeed ensure \( y \geq c_3 n^{9n d_k} \) and the above inequality holds.

Similarly using (13) of III§2 with \( c = \frac{1}{2} \) we have

\[
D_H \leq e^{O(n) d(H) 3/2} \leq e^{O(n) Q^{3/2}}
\]
and by hypothesis also

\[ n \leq \frac{1}{\alpha} \quad \text{and} \quad h \leq \lambda \]

so

\[
\frac{e^{O(n)} n/2 \log y^{8/9}}{Np} < e^{O(n)} \frac{1}{2^\alpha} \cdot \frac{5^{n/3}}{H y^{8/9}} y \frac{1/10}{1/10}
\]

\[
\leq e^{O(\ell)} y^{-1/10} \quad \text{so again if } C \text{ is large enough}
\]

then \( y \geq e^{C \ell} \) will make this term \( \ll 1 \) and we will have

\[
\sum_{\chi(H) = 1} \left| \sum_{y < N \leq u} \frac{\gamma(p) \chi(p)}{N(p)^{1+\epsilon}} \right|^2 dt \ll \frac{1}{\log y} \sum_{y < N \leq u} \frac{\left| \gamma(p) \right|^2}{Np}
\]

for \( y \geq e^{C \ell} \). Putting this into (14) leaves

\[
N_1(\alpha, T, Q; H) \ll \Delta_1 y \left( (c_{14} + 2c_{13}) r + \frac{(c_{14} + 2c_{13}) r}{(\log y)^3} \right) \cdot \left( \sum_{y < N \leq y} \frac{\left| \gamma(p) \right|^2}{Np} \right)
\]

since

\[
\int_y^{C_{12}} \left( \frac{1}{\log y} \sum_{y < N \leq u} \frac{\left| \gamma(p) \right|^2}{Np} \right) \frac{du}{u} = \frac{1}{\log y} \sum_{y < N \leq y} \frac{\left| \gamma(p) \right|^2}{Np} \int_{Np}^{C_{12}} \frac{du}{u}
\]

\[
\ll \frac{1}{\log y} \sum_{y < N \leq y} \frac{\left| \gamma(p) \right|^2}{Np} \log y.
\]
The next step then is to prove

$$\mathcal{L} \sum_{y<Np<y} \frac{|\gamma(p)|^2}{Np} \ll \Lambda_1(\log y)^3$$

(16)

which, naturally, we do by considering $e_1 = 0$ and $e_1 = 1$ separately.

Suppose first the $e_1 = 0$: then by (15) of §1 we have

$$\mathcal{L} \sum_{y<Np<y} \frac{|\gamma(p)|^2}{Np} \ll \log y \sum_{y<Np<y} \frac{\Lambda(p)^2}{Np} \leq (\log y)^2 \sum_{y<Np<y} \frac{\Lambda(p)}{N(p)}$$

$$\ll (\log y)^2 \log d_Ky \ll (\log y)^3$$

by (19) of §2 yet again, and by $d_K \leq Q \leq y$ (since $y \gg 1$ by $y \geq e^{C_2}$, and $\mathcal{L} \gg 1$, holds for large $C$). Thus (16) holds in this case, because $\Lambda_1 \gg 1$ when $e_1 = 0$.

So it remains to consider $e_1 = 1$: now we must deal directly with the function $\gamma(p)$ of lemma 11. Since the exceptional character is real-valued

$$|\gamma(p)|^2 = \left(1 + \frac{x_1(p)}{1-p_1}\right)^2 (\log Np)^2$$

where

$$\left(1 + \frac{x_1(p)}{1-p_1}\right)^2 = 2 + \frac{2x_1(p)}{1-p_1} - 1 + \frac{1}{2(1-p_1)}$$

$$= 2(1+x_1(p)) - (2x_1(p)+1) + \frac{1}{1-p_1} \left(1 - \frac{1}{1-p_1}\right)$$

$$\leq 2(1+x_1(p)) + \frac{4}{N(p)(Np-N(p))}$$
\[
\leq 2(1+x_1(p)) + 4(1-\rho_1) \log Np
\]

by the mean value theorem of elementary calculus (since \( \rho_1 < 1 \)). Thus

\[
\mathcal{L} \sum_{y < Np \leq y \cdot \zeta} \frac{|y(p)|^2}{Np} \leq 2\mathcal{L} \sum_{y < Np \leq y \cdot \zeta} \frac{(1+x_1(p))}{Np} (\log Np)^2 + 4(1-\rho_1) \mathcal{L} \sum_{y < Np \leq y \cdot \zeta} \frac{(\log Np)^3}{Np}
\]

\[
\ll (\log y)^2 \mathcal{L} \sum_{y < Np \leq y \cdot \zeta} \frac{1 + x_1(p)}{Np} + \Delta_1 (\log y)^2 \sum_{y < Np \leq y \cdot \zeta} \frac{\mathcal{L}}{Np}
\]

and since

\[
\sum_{y < Np \leq y \cdot \zeta} \frac{\mathcal{L}}{Np} \ll \log d_K y \ll \log y
\]

by (19) of \( \S 2 \) (the hypothesis of which was verified above) we are left to show

\[
\mathcal{L} \sum_{y < Np \leq y \cdot \zeta} \frac{1 + x_1(p)}{Np} \ll \Delta_1 \log y \text{ (when } e_1 = 1 \text{) (17)}
\]

But let \( E \) be the class field to \( \ker \chi_1 \): by the argument at the end of \( \S 4 \) we know that \( \rho_1 \) is the bad zero \( \beta_E \) of \( \zeta_E(s) \), that \( n_E \leq 2n \), and that \( d_E \leq q^2 \). On the other hand, if \( E \neq K \) (i.e. \( x_1 \neq 1 \)) and if \( 1 + x_1(p) \neq 0 \) then \( x_1(p) = 0,1 \) means that \( p \) is ramified, respectively split, in \( E \) ([\( E:K \)=2] so \( p \) decomposes into \( 1 + x_1(p) \) primes of \( E \) and they have norm = \( Np \). Thus
while if \( E = K \) then the same holds with a factor of 2 on the right side. In either case we can therefore apply theorem 2 to get

\[
\sum_{y < Np \leq y} \frac{1 + \chi_1(p)}{Np} \leq \sum_{y < Np \leq y} \frac{1}{N(p)}
\]

as required (since \( \beta_E = \rho_1 \)) provided that \( y \geq c_4 n_E^4 n_{E^2} \). But by the above we know that

\[
\frac{n_E^4 n_{E^2}}{c_4} \leq (16c_4)^2 n^8 n^4 \leq \frac{8\log e}{6} 4^L \leq e \leq e(\log) + \frac{8}{6} + 4L \leq 0(L)
\]

as before, and again this is \( \leq y \) provided \( y \geq cL \) for large enough \( C \). This proves (17), hence (16).

Before returning to the proof we remark that we have made no attempt to preserve optimality of \( C \) when we applied theorem 2 and 3', primarily because there are no clear criteria by which this should be done, but also because various quantities that arise do not seem to "fit" as well as might be hoped. This is also the reason for using the parameters \( \lambda, \theta \) in the statement; in particular, the application of theorem 3' entails the only occurrence of \( h_H \) as a "measure" of \( H \). Now \( h_H \) appears in theorem 3' via the corollary to lemma 8 and appears in the corollary because the error terms of lemma 8 do not combine well: one is then inclined to hope that the occurrence of \( h_H \) can be avoided altogether.
We now finish the proof of the theorem: combining (15) and (16) gives
\[ N_1(\alpha, T, Q; H) \ll \Delta_1 y \quad \text{for} \quad y \geq e^{C/2} \quad \text{with the most recent} \]
C. Putting \( y = e^{C/2} \) and recalling (5) then shows that
\[ N_1(\alpha, T, Q; H) \ll \Delta_1 e^{\sqrt{2}(c_{14} + 2c_{13})C(1-\alpha)^{1/2}} \]
a result of the quality claimed in the theorem with
\[ "c_{18}" = \sqrt{2}(c_{14} + 2c_{13})C. \]
This is however subject to (6); moreover
\[ N_1(\alpha, T, Q; H) = 0 \quad \text{for} \quad \alpha \geq 1 - 2c_{15}^{-1}, \]
by lemma 2, so it remains to consider \( \alpha \) so
\[ 1 - \frac{c_{15}^{-1}}{\sqrt{2}} < \alpha < 1 - 2c_{15}^{-1}. \] \quad (18)

Putting \( \alpha_0 = 1 - \frac{c_{15}^{-1}}{\sqrt{2}} \) the above applies and so we have for \( \alpha \) as in (18) the inequality
\[ N(\alpha, T, Q; H) \leq N(\alpha_0, T, Q; H) \ll \Delta_1 e^{c_{15}(c_{14} + 2c_{13})C} \]
from which it is clear that we can adjust the "slope" "c_{18}" to account for the remaining possibility. Thus the theorem is proved, although it is clear that incorporating (18) raises the constant c_{18} unnecessarily.

Following §6 of Bombieri [2] we show that if the exceptional zero for Q,T exists and is "sufficiently exceptional" then the zero free
region of lemma 2 can be widened by using theorem 4.

**COROLLARY.** Let \( \theta > 0 \) be given. Let \( Q \gg 1, T \gg 1 \) satisfy \( Q^T \gg 1 \) and \( Q^T \geq n^\theta \) and put \( \mathcal{L} = \log QT^n \). Suppose that the exceptional zero \((\rho_1, x_1)\) for \( Q, T \) exists and that \( \Delta_1 \leq c_{19} \). Then the region

\[
|t| \leq T, \sigma > 1 - c_{20}(\log \frac{c_{19}}{\Delta_1})\mathcal{L}^{-1}, d_x \leq Q (x \text{ primitive})
\]

contains at most the zero \((\rho_1, x_1)\). Here \( c_{19}, c_{20} \) depend only on \( \theta \).

**PROOF:** We will apply theorem 4 with a fixed \( \lambda \) which we now construct:

let \( C \) be the constant of (9) of III\( \S 2 \) so that

\[
h_H \leq e^{Cm}d_xN(m)
\]

and let \( C' \) be the constant so that \( \mathcal{L} \geq C'n \) (which comes from the condition \( Q^T \gg 1 \) of theorem 4). We put \( \lambda = CC' + 1 \) and fix this value.

suppose now that \( (\rho, x) \), \( \rho = \beta + iy \) is a zero with \( |y| \leq T, d_x \leq Q \): we must show if this zero is non-exceptional then

\[
\beta \leq 1 - c_{20}(\log \frac{c_{19}}{\Delta_1})\mathcal{L}^{-1}
\]

we let \( c_{19}, c_{20} \) be determined by theorem 4 for \( \lambda \) (fixed) and \( \theta \) so that \( c_{19}^{-1} \) is the "c_{18}" of theorem 4 and \( c_{19}^{-1} \) is the constant of theorem 4 implied by \( \ll \). Suppose \( (\rho, x) \) above is non-exceptional and put \( H = \ker x \) : then \( \delta_H = \delta_x \) hence \( d(H) = d_x \leq Q \) and

\[
h_H \leq e^{Cm}d_xN(\delta_x) = e^{Cm}d_x \leq e^{CC'\mathcal{L}}Q \leq e^{(CC'+1)\mathcal{L}} = e^{c\mathcal{L}}
\]
and by theorem 4 we have

\[ 1 \leq N_1(\beta, T; Q; \chi) \leq N_1(\beta, T; Q; \mathcal{H}) \leq c_{20}^{-1} (1-\beta)^{2L} \]

from which (19) follows.

Note that \( h_H \) has been successfully eliminated by (9) of III§2: this would work in other cases, however, apparently, not in general (at least not unless we want to replace \( d(H) \) by \( d_K N(h_H) \) as the basic measure). Finally we point out that, again as in Bombieri [2], theorem 4 implies Siegel's theorem (which gives a sharper lower bound for \( 1 - \rho_1 \) than theorem 1') however Siegel's theorem is not effective, even for \( K = \mathbb{Q} \), and the \( K \)-uniformity above will not help in any way.
Finally we are prepared to obtain explicit and effective bounds for the least prime ideal in a generalized arithmetic progression. Actually, after Fogels [8], we do rather more and prove the existence of prime ideals in "short" intervals. Our approach, however, differs from that of Fogels technically, if not in spirit, in that we find that the "averaging functions" of chapter III make the estimations somewhat easier. The proof of the existence of these prime ideals is, in outline, obtained by first establishing an "explicit formula" (which will be the analogue, for prime ideals, of lemma 8) which makes the solution depend on the distribution of zeros of L-series, and then applying theorem 4. After proving this existence theorem we then apply it to the problem of the least prime ideal in "Tchebotarev classes".

§1 An explicit formula

Once again we fix our field \( K \); we begin by proving the formula of the title, which is not the "usual" explicit formula of prime number (ideal) theory for reasons which will be apparent later. There are effective versions of the "usual" explicit formula available in Lagarias and Odlyzko [16], for example. However the version we will use is based on the functions \( \eta_A \) with parameter \( A > 0 \) of III§1. Employing the notation used there we can state

**Lemma 12.** Let \( a \) be such that \( 0 \leq a \leq \frac{1}{3} \), and suppose that the integer
\[ l \geq 3, \text{ and that the parameter } A \text{ satisfies } A \geq l, A > 1. \text{ Then for } y \geq d_k, y > 1 \text{ we have}
\]
\[ \sum_{\rho \in \mathcal{L}(x)} h_\chi(\rho - 1)y^{\rho - 1} + O(y^{-a}T \log dT^n) + \frac{1}{T^2} \log dT^n \]

for any \( T \geq A \). Here \( \chi \) is any primitive congruence class character of \( K \).

**Proof:** We begin by observing that
\[ \sum_{p} \frac{x(p) \log N(p)}{N(p)} h_\chi\left(\frac{y}{Np}\right) = \sum_{a} \frac{\Lambda(a) \chi(a)}{Na} h_\chi\left(\frac{y}{Na}\right) + (Ay^{-1/2} \log y) \]  

for since \( \Lambda(p) = \log Np \) and \( \Lambda(a) = 0 \) if \( a \) is not a prime power the difference between our sums is in absolute value

\[ \leq \sum_{a=p^m, m \geq 2} \frac{\Lambda(a) \chi(a)}{Na} h_\chi\left(\frac{y}{Na}\right) \leq \frac{\Lambda}{2} \sum_{p, m \geq 2} \frac{\log N(p)}{N(p)^m} \]

by (2), (10) of III§1. Now \( e^{y/A} \leq 4 \) (by \( \lambda \leq A \)) and

\[ \sum_{m \geq 2} \frac{1}{N(p)^m} = \frac{1}{N(p)^2 - 1} \ll \frac{1}{N(p)^2} \]

\( \ll 1 \) so our sum is

\[ A \sum_{\frac{1}{4} y \leq Np^2 \leq 4y} \frac{\log Np}{Np^2} = A \sum_{\frac{1}{2} y \leq Np \leq 2y} \frac{1}{Np} \cdot \frac{\log Np}{Np} \ll A y^{-1/2} \sum_{Np \geq 2y^{1/2}} \frac{\log Np}{Np} \]
\( A y^{-1/2} \log y \) by (19) of I§2 (since \( d \geq y \) and \( y \gg 1 \)).

We now proceed as in the proof of lemma 8 starting from

\[
\sum_{a} \frac{\Lambda(a) \chi(a)}{N(a)} H_{\alpha} \left( \frac{y}{Na} \right) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} -\frac{L'}{L}(s+1, \chi) h_{\alpha}(s) y^s ds
\]

for by lemma 7 the left side is

\[
\sum_{a} \frac{\Lambda(a) \chi(a)}{N(a)} \frac{1}{2\pi i} \int_{-\infty}^{+\infty} h_{\alpha}(s) y^s ds = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \left( \sum_{a} \frac{\Lambda(a) \chi(a)}{N(a)^{s+1}} \right) h_{\alpha}(s) y^s ds
\]

which, by (16) of I§2, is the right side.

The first step is to move the line of integration from \( \sigma = 1/2 \) to \( \sigma = -3/2 \) so, as before, we consider the positively oriented rectangle \( R_U \) with vertices at \(-3/2 - iU, 1/2 + iU, 1/2 - iU, -3/2 + iU\) where \( U > 0 \) is large (we let \( U \to \infty \)), and will be more precisely handled below. Clearly we need some estimates for \(-\frac{L'}{L}(s, \chi)\) on the boundary of our rectangle, which we now proceed to develop.

First if \( m \) is an integer \( \geq 1 \) then the number of zeros \( \rho \) of \( L(s, \chi) \) which satisfy

\[
|\text{Im} \rho - m| < 1 \text{ or } |\text{Im} \rho + m| < 1
\]

is \( \ll \log d \chi^m \) by the corollary to lemma 4 (with \( Q = d \chi \), \( T = m \) and \( \tau = \pm m \)). Thus there exists a \( U = U_m \) with \( |U_m - m| < 1 \) so that every zero \( \rho \) of \( L(s, \chi) \) satisfies

\[
|\text{Im} \rho - U| \gg (\log d \chi^m)^{-1} \text{ and } |\text{Im} \rho + U| \gg (\log d \chi^m)^{-1}
\]
for \( U = U_m \). Such values of \( U \) we call admissible and we observe that we can choose a sequence of admissible values tending to \( \infty \).

The estimation on horizontal edges of \( R_U \) will then come from

\[
\left| \frac{L'(s \pm iU, \chi)}{L(s, \chi)} \right| \ll (\log d_n)^2 \quad \text{uniformly for } \frac{1}{2} \leq \sigma \leq \frac{3}{2} \tag{3}
\]

if \( U \) is admissible and \( \gg 1 \). For \( \frac{1}{2} \leq \sigma \leq \frac{3}{2} \) this follows from lemma 4 (with \( Q = \chi, T = U \) and \( \tau = \pm U \)) since, putting \( s = \sigma \pm iU \) there gives

\[
\frac{L'(s, \chi)}{L(s, \chi)} = \frac{\delta_\chi}{s-1} - \sum_{\rho \in \mathcal{L}(\chi)} \frac{1}{s-\rho} + O(\log d_n^2)
\]

where the \( \delta_\chi \) term is \( O(1) \) by \( U \gg 1 \) and if \( \rho \in \mathcal{L}(\chi) \) then

\[
|s-\rho| \geq |\pm U - \Im \rho| \gg (\log d_n)^{-1}
\]

so that

\[
\left| \sum_{|\rho-(1+iU)|<1} \frac{1}{s-\rho} \right| \ll \log d_n^2 \sum_{|\rho-(1+iU)|<1} 1 \ll (\log d_n^2)^2
\]

by the corollary to lemma 4 again. To prove (3) in the remaining case \( -\frac{1}{2} \leq \sigma \leq \frac{1}{2} \) we use the functional equation: namely by (4), (5) of \( \text{I} \& \text{I} \) we have

\[
L(s, \chi) = \frac{\Gamma_X(s-1)}{\Gamma_X(s)} W(\chi)^{-1} \frac{A_X^{1-s} \Gamma_X(1-s)}{A_X^{s} \Gamma_X(s)} L(1-s, \overline{\chi})
\]

since \( A = A_X \), \( \Gamma = \Gamma_X \); expanding \( \frac{\Gamma_X(s)}{\Gamma_X(1-s)} \) by (4) of \( \text{I}\&\text{I} \), and taking logarithmic derivatives leaves
\[
\frac{L'}{L} (s, \chi) = n \log 2\pi - \log d_\chi + (r_1 - \mu_\chi) \frac{\pi}{2} \cot \frac{\pi}{2} s
\]

\[+ \mu_\chi \frac{\pi}{2} \cot \frac{\pi}{2} (s+1) + r_2 \pi \cot \pi s
\]

\[- n \Gamma' \frac{(1-s)}{-} - \frac{L'}{L} (1-s, \chi) .
\]

However, from \( \cot \pi s = \frac{e^{2\pi i s} + 1}{e^{2\pi i s} - 1} = \frac{1 + e^{-2\pi i s}}{1 - e^{-2\pi i s}} \), we see that, on writing \( s = \sigma + it \), we have

\[
\cot \pi s \to -i \text{ as } t \to +\infty \text{ and }
\]

\[
\cot \pi s \to i \text{ as } t \to -\infty
\]

and that \( \cot \pi s \) has poles only at the integers, so in particular \( \cot \pi s \) is bounded in the vertical strip \( -\frac{1}{2} \leq \sigma \leq \frac{3}{2} \) outside of the neighborhoods \( |s| \leq \frac{1}{4} \), \( |s-1| \leq \frac{1}{4} \) of its poles there. Thus we have

\[
\frac{-L'}{L} (s, \chi) = \log d_\chi + n \Gamma' \frac{(1-s)}{-} + \frac{L'}{L} (1-s, \chi) + o(n) \tag{4}
\]

for \( -\frac{1}{2} \leq \sigma \leq \frac{3}{2} \), \( |s| \geq \frac{1}{4} \), \( |s-1| \geq \frac{1}{4} \); since \( \mu_\chi, r_1, r_2 \) are all \( \leq n \). Note also that the boundedness argument for \( \cot \pi s \) above applies to the other cotangent terms (which have poles only at even and odd integers, respectively) and that it is easy, if tedious, to get these bounds explicitly.

We can now finish the outstanding case \( -\frac{1}{2} \leq \sigma \leq \frac{1}{2} \) of (3): for in (4) the \( \frac{L'}{L} (1-s, \chi) \) term is now \( \ll (\log d_\chi v^n)^2 \), because \( 1 - s \) falls into the case \( \frac{1}{2} \leq \sigma \leq \frac{3}{2} \) already handled above (with \( \chi \) replaced
by  \( \chi \) , so it remains only to verify that

\[
\frac{\Gamma'(1-s)}{\Gamma(1-s)} \ll \log U^n \quad \text{for } U > 1
\]

which is immediate from

\[
\frac{\Gamma'(s)}{\Gamma(s)} = \log s + O(\frac{1}{|s|}) \quad \text{uniformly in } \sigma \geq \frac{1}{2}
\]

(since 1-s again falls into this region and U large allows us to absorb \( \arg s \)) which was discussed in the proof of (1) of III§3. Therefore (3) is proved and we can turn to

\[
\left| \frac{L'}{L}(-\frac{1}{2}+it, \chi) \right| \ll \begin{cases} 
\log d_x + O(n), & |t| \leq 3 \\
\log d_x |t|^n, & |t| \geq 3
\end{cases}
\]

for which we can again apply (4), observing that \( \left| \frac{L'}{L}(\frac{3}{2}-it, \chi) \right| \ll n \), which is proved in the same way as (13) of III§3. Thus (4) and (5) yield

\[
-\frac{L'}{L}(-\frac{1}{2}+it, \chi) = \log d_x + n \log(\frac{3}{2}-it) + O(n)
\]

and (6) follows, since when \( |t| \geq 3 \) (i.e. is "large") we can absorb the \( O(n) \) term in \( \log |t|^n \).

We now can return to the integral (2) and our rectangle \( R_U \) with \( U \to \infty \) through admissible values: we first estimate the integrals on the horizontal edges, qualitatively, by (3) and by (12) of III§1, and they are, in absolute value
Moreover the integrand of the integral (2) has (inside our rectangle \( R_U \)) a pole at \( \rho - 1 \) with residue \(-h_\lambda(\rho-1)y^{\rho-1}\) for each non-trivial zero \( \rho \) of \( L(s,\chi) \) (and this term occurs as often as the multiplicity of \( \rho \)). In addition there is a pole at \( s = -1 \) (usually) corresponding to the trivial zero of \( L(s,\chi) \) at 0; the residue is \(-h_\lambda(-1)y^{-1}\).

Finally in case \( \delta_X = 1 \) there is a pole at \( s = 0 \) corresponding to the simple pole of \( \zeta_k(s) \) at \( s = 1 \); the residue is \( h_\lambda(0)y^0 = 1 \). Putting this all together and letting \( U \to \infty \) through admissible values leaves us with

\[
\frac{1}{2\pi i} \int_{\frac{1}{2} - i\infty}^{\frac{1}{2} + i\infty} \left( \log \frac{U^n}{\sigma} \right)^2 \frac{A^\lambda}{|\sigma + 1 \pm iU|^{\frac{3}{2}}} e^{A^{-1}\lambda |\sigma|} \sigma d\sigma \ll \frac{1}{2} \int_{-\frac{3}{2}}^{\frac{3}{2}} (\log U^n)^2 d\sigma
\]

\[
\ll U^{-\lambda}(\log U)^2 \quad \text{which tends to} \quad 0 \quad \text{as} \quad U \to \infty.
\]

where the \( O \)-term comes from the trivial zero mentioned above: more precisely we note that

\[
|h_\lambda(-1)| \leq \left(1 + \frac{1}{5A^2}\right)^{\lambda} \leq \left(1 + \frac{1}{5\lambda^2}\right)^{\lambda} \ll 1
\]
by (13) of III§1 and $A \geq \xi$, and that the multiplicity of the trivial zero is $\leq n$ by the discussion preceding (10) of I§1 (since there are $\leq n$ poles from $\Gamma$-factors to cancel).

Actually the statement is not yet completely proved: in place of $\sum_{\rho} + \int_{-\frac{3}{2} - i\infty}^{-\frac{3}{2} + i\infty}$ we only know that $\lim_{\text{U} \to +\infty} \left| \text{Im } \rho \right| < \text{U}$ admissible $\int_{-\frac{3}{2} - i\infty}^{-\frac{3}{2} + i\infty}$ can be put, since we have not yet verified that the sum and integral converge independently. This (and more) is provided by

$$\int_{-\frac{3}{2} + i\infty}^{-\frac{3}{2} - i\infty} \left| \frac{-L}{L}(-s+1, x)h_x(s)y^s \right| ds \ll y^{\frac{3}{2}} \log d_x A^n$$

(8)

for the integral is

$$\leq y^{\frac{3}{2}} \left( \int_{|t| \leq A} \left| \frac{-L}{L}(-\frac{1}{2} + it, x) \right| h_x(-\frac{3}{2} + it) dt \right)$$

where, for $|t| \leq A$, we can combine the two cases of (6) (because of $A \ll 1$) to give $\left| \frac{L}{L}(-\frac{1}{2} + it, x) \right| \ll \log d_x A^n$, when the boundedness of $h_x(-\frac{3}{2} + it)$, which follows from (15) of III§1 (by $A \geq \xi$ and $\xi \geq 3$), shows that $\int_{-A}^{A} \left| \frac{L}{L}(-\frac{1}{2} + it, x)h_x(-\frac{3}{2} + it) dt \right| d t \ll A \log d_x A^n$.

On the other hand using (6) and (12) of III§1 the integral over $|t| \geq A$ is
\begin{align*}
&\ll \int_{A}^{\infty} \log d \cdot |t|^n \frac{A^{-\frac{t}{2}}}{|t|^{-\frac{3}{2} + it|}} \, dt \\
&\ll A \int_{A}^{\infty} \frac{\log d + n \log t}{t^{\frac{1}{2}}} \, dt
\end{align*}

since the upper bounds are even functions of \( t \) and since \( A^{-\frac{t}{2}} \leq 1 \).

Finally, integration by parts shows that the last integral is equal to

\begin{align*}
&\frac{A^{-\frac{t}{2}}(\log d + n \log A)}{t - 1} + \int_{A}^{\infty} \frac{n}{t^{\frac{1}{2}}} \frac{1}{t} \, dt \\
&\leq A^{-\frac{t}{2}} \log d \cdot A^n + nA^{-\frac{t}{2}} \text{ which, by } A \gg 1, \text{ proves (8).}
\end{align*}

Finally we turn to the \( \sum_{\rho \in \mathcal{L}_{\chi}(\chi)} \) sum of (7), and begin by proving

\begin{equation}
\left| \sum_{\substack{\rho \leq T \Re \rho \leq \frac{1}{2} - \frac{1}{\log y} \log T \log d}} |h_{\frac{t}{2}}(\rho - 1)y^{\rho - 1}| \ll y^{-\frac{a}{2}} T \log d \cdot T^n
\end{equation}

for which we begin by applying (15) of III\&1 (which applies because

\begin{align*}
&|\Re(\rho - 1)| \leq \frac{1}{2} \log T \log d \text{ and } |y^{\rho - 1}| \leq y^{-\frac{a}{2}} \text{ to conclude that}
\end{align*}

\begin{align*}
&\sum_{\substack{\rho \leq T \Re \rho \leq \frac{1}{2} - \frac{1}{\log y} \log T \log d}} |h_{\frac{t}{2}}(\rho - 1)y^{\rho - 1}| \ll y^{-\frac{a}{2}} \sum_{\substack{\rho \leq T \Re \rho \leq \frac{1}{2} - \frac{1}{\log y} \log T \log d}} 1
\end{align*}

from which (9) follows by

\begin{equation}
\sum_{\rho \in \mathcal{L}_{\chi}(\chi), \left| \Im \rho \right| \leq T} 1 \ll T \log d \cdot T^n \text{ for } T \gg 1
\end{equation}

This follows from the corollary to lemma 4 as already observed in the
paragraph following it.

The other part of the $\sum_{\rho}$ sum we must estimate is done by

$$\sum_{|\text{Im } \rho| \geq T} |h^\rho (\rho - 1)y^{\rho - 1}| \ll \frac{A^\rho}{T^\rho} T \log d \chi^T$$

the proof of which begins with (12) of III§1 and $|y^{\rho - 1}| \leq 1$ (by $\text{Re } \rho \leq 1$) from which we deduce that our sum is

$$\ll \sum_{|\text{Im } \rho| \geq T} \frac{A^\rho}{|\rho - 1|^\rho} e^{A^{-1} \rho \text{Re } \rho - 1} \ll A^\rho \sum_{|\text{Im } \rho| \geq T} \frac{1}{|\text{Im } \rho|^\rho}$$

since $A^{-1} \rho \text{Re } \rho - 1 \leq A^{-1} \rho \leq 1$. To estimate the remaining sum we define, for each $\rho \in \mathcal{L}(\chi)$ with $|\text{Im } \rho| \geq T$, a counting function $f_\rho$ by

$$f_\rho (t) = \begin{cases} 1 & t \geq |\text{Im } \rho| \\ 0 & |\text{Im } \rho| > t \geq T \end{cases}$$

and observe that by (10) we have

$$\sum_{\rho} f_\rho (t) = \text{the number of } \rho \text{ with } T \leq |\text{Im } \rho| \leq t$$

$$\ll t \log d \chi^T \text{ (for } t \geq T \gg 1)$$

while clearly

$$\int_T^\infty \frac{\ell f_\rho (t)}{t^{\ell + 1}} \, dt = \int_{|\text{Im } \rho|_T}^\infty \frac{\ell}{t^{\ell + 1}} \, dt = \frac{1}{|\text{Im } \rho|^{\ell}}.$$

Then
\[
\sum_{|\text{Im } \rho| \geq T} |\text{Im } \rho|^{-\frac{\gamma}{2}} = \int_{T}^{\infty} \frac{\frac{\gamma}{2}+1}{t} \sum_{\rho} f_{\rho}(t) \, dt \ll \int_{T}^{\infty} \frac{\log d_{\chi} + n \log t}{t^\frac{\gamma}{2}} \, dt
\]

\[
\ll \frac{\gamma}{2-1} T^{1-\frac{\gamma}{2}} \log d_{\chi} T^{n} \text{ by the last step in the proof of (8). Thus (11) is proved, keeping in mind that the condition } T > 1 \text{ used above can be met by } T \geq \Lambda .
\]

Using successively the relations (1), (2), (7), (8), (9), (11) we find that

\[
\sum_{\rho} \frac{\chi(\rho)}{N_{\rho}} \log \frac{N_{\rho}}{N_{p}} H_{\chi}(\frac{y}{N_{\rho}}) = O(\gamma^{-1/2} A \log y) + O(ny^{-1})
\]

\[+ O(y^{-3/2} A \log d_{\chi} A^{n}) + O(y^{-a} T \log d_{\chi} T^{n}) + O(\frac{A^{\gamma}}{T^{\gamma}} T \log d_{\chi} T^{n})
\]

\[+ \delta_{\chi} - \sum_{\rho \in \mathbb{C} \chi \rho} \frac{h_{\chi}(\mu-1) \gamma^{-1}}{|\text{Im } \rho| < T, \text{Re } \rho > 1-a}
\]

and it remains only to absorb the first three error terms into the fourth: but for \( y \gg 1 \): \( y^{-1/2} A \log y \ll y^{-1/3} A \leq y^{-a} T \) (by \( a \leq \frac{1}{3} \), \( \Lambda \leq T \)); and by the Minkowski bound for discriminants we have \( ny^{-1} \ll d_{\chi}^{2/3} y^{-1} \leq y^{-1/3} \) by \( d_{\chi} \leq y \); finally the third error term can be absorbed by \( a \leq 3/2 \) and \( \Lambda \leq T \), proving lemma 12.

The advantage of this explicit formula over the "usual" one is in the error term \( O(\frac{A^{\gamma}}{T^{\gamma}} T \log d_{\chi} T^{n}) \) which for \( T \) somewhat larger than \( A \) (and suitable \( \chi \)) is still "small enough" (the analogue for the "usual" explicit formula is a term of size about \( 1/T \)). We also prefer lemma 12 to the similar formula of Fogels [8] (which replaces our \( \eta_{\chi} \) with the function \( \eta(x) = \frac{1}{2\sqrt{x}A} e^{-x^2/2A} \) where \( A \) is again a parameter) because now \( H_{\chi} \) already vanishes outside an interval of suitable size, which saves the effort of showing that the outside terms are insignificant
Finally we point out that the argument of (1) would allow us to replace the sum of the lemma by

$$\sum_{p, \deg p=1} x(p) \frac{\log N_p}{N_p} H_N(y)$$

where by \( \deg p \) we mean the degree of the residue class field at \( p \) over the prime field. Thus the later results can be viewed as asserting the existence of primes \( p \) with \( \deg p = 1 \), and satisfying certain (more significant) conditions. Since we will have occasion to use this remark we include a proof: by (1) it suffices to show that

$$\sum_{\deg p \geq 2} x(p) \frac{\log N(p)}{N_p} H_N(y)$$

is small enough that it can be absorbed in the error terms of lemma 12. The first step in the proof of (1) (i.e. using the properties of \( H_N \)) applies here and shows our sum is in absolute value

$$\ll A \sum_{\frac{1}{4}y \leq Np \leq 4y} \frac{\log Np}{Np} \leq n A \sum_{\frac{1}{4}y \leq p \leq 4y} \frac{\log p^m}{p^m}$$

for some \( m \geq 2 \) since there are \( \leq n \) prime ideals of \( K \) lying above a single prime \( p \) of \( \mathbb{Q} \). Since every \( p \) is \( \geq 2 \), this is

$$\leq n A \sum_{2 \leq m \leq \frac{\log 4y}{\log 2}} \frac{1}{m} \sum_{\frac{y}{4^m} \leq p \leq \frac{y}{m}} \frac{1}{p^m} \ll n A \sum_{2 \leq m \leq \log y} \frac{1}{y} \sum_{p \leq (4y)^m} \frac{1}{p^m}$$

$$= n A^{-1} \left( \sum_{m \leq \log y} \left( \sum_{p \leq 2y^2} \frac{1}{p} \log p \right) \right) \ll n A^{-1} (\log y)^2 y^2$$
by the prime number theorem (for $Q$, hence $K$ uniform). But $n \ll d_{k}^{1/12} \leq y^{1/12}$ by the Minkowski bound, and $(\log y)^2 \ll y^{1/12}$, $A \leq T$ (by the hypotheses of lemma 12) then show the above sum is $\ll y^{-1/3}T$ which is smaller than the first error term of lemma 12, proving our contention.

§2 Prime ideals in short segments of arithmetic progressions

In this section we arrive at the arithmetic goal of this work. For convenience we first prove an awkward version of the existence theorem; nevertheless our strongest result will be

**Lemma 13.** Let $\lambda > 0, \theta > 4$ be given, and let $H$ be a primitive congruence class group (in the sense of III§2) and $C$ a coset mod $H$.

Suppose the parameter $A$ satisfies

$$A \geq d(H)^{\frac{1}{2n}}, A \geq \frac{\theta}{4}, A \geq h_{H}^{\frac{1}{4\lambda n}}, A \gg 1.$$ 

Then for all $y \geq A^{c_{21}n}$ we have

$$h_{H} \sum_{\substack{p \in C \\ \log Np \gg \min(c_{K}^{-1}A^{-1}, A^{-3})}} \frac{\log Np}{Np} \gg \min(c_{K}^{-1}A^{-1}, A^{-3})$$

$$\gg N(p) < ye^{c_{22}n/A} < N(p) < ye^{c_{22}n/A}$$

where the constants $c_{21}, c_{22}$ and that implied by $\gg$ depend only on $\lambda, \theta$ and where $c_{K}$ is defined in the corollary of I§4 (and depends at most on $n$).
PROOF: As the notation suggests this result is a consequence of lemma 12 and theorem 4; in fact somewhat more can be proved by optimizing more carefully, however since we have, as yet, little idea of the size of our constants (except the guess that they are unreasonably large) we do not pursue this.

We begin by applying lemma 12 to the characters $\chi$ with $\chi(H) = 1$ and with

$$T = A^2, \quad a = 1 - c_{17} \quad \text{(of theorem 4 for this } \lambda, \theta \}$$

and $\ell$ is the smallest integer $\geq \left(\frac{8}{\theta} + 4\lambda + 6\right)n$.

Note that the conditions of lemma 12 for $T, A$ are satisfied (in particular $A \geq \ell$ is satisfied because $\ell \ll n$ and $A \geq n^2$ with $\theta > 4$ shows it holds for large $n$, while small $n$ can be accomodated by $A \gg 1$).

Before we can use lemma 12 we must make the transition from characters of $H \mod \delta_H$ to primitive characters (as in lemma 12), in the terminology of III§2. But if $Np \leq y e^{-2/A}$ then $H_{y/Np}(y/Np) = 0$ (by (10) of III§1) so, since $\ell/A \leq 1$, if we arrange that every $p/H$ has $Np \leq y e^{-1}$ then the primitive characters and characters $\mod \delta_H$ will agree throughout the range of our sum (as in the proof of (14) of IV§2). Again, since $H$ is primitive, $p/H$ implies $p/\delta_\psi$ for some $\psi$ with $\psi(H) = 1$ (by (10) of III§2) hence $N(p) \leq N(\delta_\psi) \leq d(H) \leq A^{2n}$. Thus if $y \geq eA^{2n}$ then lemma 12 applies to $\chi$ with $\chi(H) = 1$; in particular if we insist that
then $c_{17} > 9/10$ (see first paragraph of proof of theorem 4) and $\epsilon \geq 6n$ implies $y \geq A^{60n}$ which certainly is enough. Moreover now $y \geq d(H)^{30}$ $\geq d_{K}^{30}$ so the hypothesis of lemma 12 on $y$ are also satisfied. Thus applying lemma 12 and using $d_{x} \leq d(H) \leq A^{2n}$ for $x$ with $x(H) = 1$ gives

$$
(1-c_{17})^{-1} y \geq A \quad (2)
$$

and the reason for (2) is that it implies that $y \leq A^{-\epsilon}$ so the first error term can be absorbed in the second. We take a closer look at the remaining error term, and use $\log A^{4n} \ll \log A^{n} \ll A^{2n}$, and the restriction on $A$ made in the statement of the lemma, to see that

$$
\frac{8}{9} \frac{(4\lambda + 2 + \frac{1}{2})n}{2} - \frac{8}{9} + 4\lambda + 6)n \leq \frac{1}{\epsilon} \leq (4A^{-2}) A^{-3n} \leq A^{-3} \quad \text{for } A \geq 16 \quad (i.e. A \gg 1)
$$

allowing us to write

$$
\begin{align*}
&h_{H} \sum_{p \in \mathbb{C}} \frac{\log Np}{np} H_{\chi}(\frac{y}{np}) = 1 - \sum_{\chi(H)=1} \bar{\chi}(C) \sum_{\rho \in \mathbb{C}^{*}} \bar{\chi}(\rho) \frac{\log Np}{N(p)} H_{\chi}(\frac{y}{np}) \\
&\quad + O(h_{H} y^{\frac{1}{2}} A^{2} \log A^{4n}) + O(h_{H} A^{-\epsilon} A^{2} \log A^{4n})
\end{align*}
$$

(4)
Before estimating the sum \( \sum_{\rho} \) we observe that if we put
\[
c_{22} = \frac{8}{6} + 4\lambda + 7
\]
then we have \( \lambda \leq c_{22} \) and by (10) of III§1 we deduce immediately that
\[
\frac{A}{2} \sum_{\rho \in \mathcal{C}} \log \frac{N\rho}{Np} \leq \frac{c_{22} n}{A} \frac{c_{22} n}{A} < Np < ye
\]
so only the \( \sum_{\rho} \) sum stands as an obstacle to the lemma. But this is precisely what theorem 4 prepares us for; we put
\[
Q = A^{2n} , \text{ hence } Q = T^n
\]
when \( QT^n = A^n \), \( \frac{1}{Q} T = A \) show that the conditions put on \( A \) are precisely those needed for theorem 4. However theorem 4 applies only to non-exceptional zeros (for \( Q,T \)) so in addition to the notations \( \mathcal{L}(\chi), \mathcal{L}, \Delta_1, e_1, (\rho_\lambda, \chi_\lambda) \) of IV§2 we define
\[
e_1(H) = \begin{cases} 1, & \text{if } e_1 = 1 \text{ and } \chi_1(H) = 1 \\ 0, & \text{if not} \end{cases}, \quad \Delta_1(H) = \begin{cases} \Delta_1 & \text{if } e_1(H) = 1 \\ 2c_1 & \text{if not} \end{cases}
\]
where \( \chi_1(H) = 1 \) is an abuse of notation which we intend to mean: the character mod \( \mathcal{H} \) induced by \( \chi_1 \) is trivial on \( H \), if indeed \( \chi_1 \)
induces a character mod \( h_H \) (i.e. if \( h_{\chi^1} \) divides \( h_H \)). Of course with \( Q = A^{2n} \), \( T = A^2 \) we have

\[
\mathcal{L} = \log A^{4n}
\]  

and taking the \((\rho^1, \chi^1)\) term out of the \( \sum_\rho \) sum in (4) we can now write

\[
\frac{1}{2A} h_H \sum_{\rho \in \mathcal{L}} \frac{\log N\rho}{N\rho} \geq (1-e_{h_H}(\chi^1)(c)h_{\chi^1}(\rho^1-1)\gamma^\rho_1-1) \\
- c_2^{2n/A} ye^{c_2^{2n/A}} < N\rho < ye^{c_2^{2n/A}}
\]

\[
\frac{1}{2A} h_H \sum_{\chi(H)=1} \sum_{\rho \in \mathcal{L}_2(\chi)} h_{\chi^1}(\rho^1-1)\gamma^\rho_1 + \mathcal{O}\left(\frac{1}{(2n)!A^3}\right)
\]

and are now ready to tackle the \( \sum_\rho \) sum. Indeed

\[
| \sum_{\chi(H)=1} \sum_{\rho \in \mathcal{L}_2(\chi)} h_{\chi^1}(\rho^1-1)\gamma^\rho_1 | \ll \Delta_1 \exp\left(-2c_1 \frac{\log \gamma}{\mathcal{L}}\right),
\]

for which we define, for non-exceptional zeros \( \rho \) appearing in the sum (i.e. \( L(\rho, \chi) = 0 \) for \( \chi \) with \( \chi(H) = 1 \)); and \( |\text{Im} \ \rho| < A^2, \text{Re} \ \rho > c_{17} \), a counting function \( g_{\rho} \) by

\[
g_{\rho}(\alpha) = \begin{cases} 
0 & 1 \geq \alpha > \text{Re} \\
1 & \text{Re} \ \rho \geq \alpha > c_{17}
\end{cases}
\]

Since every such \( \rho \) satisfies \( \text{Re} \ \rho < 1 - 2c_1 \frac{\mathcal{L}}{\rho^1} \) (i.e. is not the exception to lemma 2) we have
and moreover

\[ \sum_{\chi(H)=1} \sum_{\rho \in \mathcal{L}_1(x)} \frac{g_\rho(a)}{|\text{Re } \rho| < A^2, \text{Re } \rho > c_{17}} \]

of non-exceptional zeros \( \rho \) of \( L(s, \chi) \) with \( \chi(H) = 1 \), satisfying

\[ |\text{Im } \rho| < A^2, \text{Re } \rho \geq \alpha, = N_1(\alpha, T, Q; H) \text{ of theorem } 4. \]

Now by (15) of III§1, our sum (10) is certainly

\[ \ll \sum_{\chi(H)=1} \sum_{\rho \in \mathcal{L}_1(x)} y \text{Re } \rho^{-1} = \frac{1}{y} \sum_{\rho} y \text{Re } \rho \]

\[ = \frac{1}{y} \sum_{\rho} (y^{c_{17}} + \log y^{1-2c_{17}^{\nu-1}}) \int_{c_{17}} y^\alpha g_\rho(\alpha) d\alpha \]

\[ = \frac{1}{y} y^{c_{17}} N_1(c_{17}, Q, T; H) + \log y^{1-2c_{17}^{\nu-1}} \int_{c_{17}} y^\alpha N_1(\alpha, Q, T; H) d\alpha \]

\[ \ll \frac{\Delta_1}{y} y^{c_{17}} (A^4 n)^{c_{18}(1-c_{17})} + \log y^{1-2c_{17}^{\nu-1}} y^{\alpha (A^4 n)^{c_{18}(1-\alpha)}} d\alpha \]

by theorem 4. But this integral equals
\[
\frac{4c_{18}^n}{A} \int_{c_{17}}^{1-2c_{18}^{\Delta_{1}}-4c_{18}^n} (yA_{18}^{1-4c_{18}^n}) \, da = \frac{4c_{18}^n}{A} \left\{ -4c_{18}^n \left( 1-2c_{18}^{\Delta_{1}} -4c_{18}^n \right) \right\} \frac{1-2c_{18}^{\Delta_{1}} -4c_{18}^n}{\log(yA_{18}^{1-4c_{18}^n})}
\]

\[
= \frac{y}{\log(yA_{18}^{1-4c_{18}^n})} \left( -2c_{18}^{\Delta_{1}} -4c_{18}^n \right) \frac{8c_{18}^{\Delta_{1}} -4c_{18}^n}{\log(yA_{18}^{1-4c_{18}^n})} - \frac{c_{17} (A_{4n}^{c_{18} (1-c_{17})})}{\log(yA_{18}^{1-4c_{18}^n})}
\]

so, since \( A_{4n}^{c_{18} (1-c_{17})} = e \), and since \( \frac{\log y}{-4c_{18}^n} > 1 \) enables this negative term to cancel the first term above, we have found our sum (10) is

\[
\ll \frac{\Delta_{1}}{y} \frac{y \log y}{-4c_{18}^n} e^{-2c_{18}^{\Delta_{1}} \log y} e^{2c_{18}^{c_{18}}} \ll \Delta_{1} e^{-2c_{18}^{\Delta_{1}} \log y}
\]

provided that

\[
y \geq A_{18}^{8c_{18}^n}
\]

(11)

since then \( \frac{\log y}{-4c_{18}^n} \leq 2 \) will hold.

So we must show that the first term on the right side of the inequality (9) is "large" which is done by

\[
1 - e_{1}(H)X_{1}(C)h_{e}(\rho_{1}-1)y^{\rho_{1}-1} \gg \Delta_{1}(H)
\]

(12)
which makes sense since $\rho_1$ is real and $\chi_1$ real valued (being the exception to lemma 2). Indeed $3/4 < \rho_1 < 1$ holds, by IV§1. Since $h_\chi$ takes values $\geq 1$ for real arguments (as in proof of (15) of III§2), it is clear that there is nothing to prove unless $e_1(H) = 1$ and $\chi_1(C) = 1$, when we must show $1 - h_\chi(\rho_1^{-1})y^{\rho_1^{-1}} \gg \Delta_1$. We write

$$\phi = \chi_1^{-1} \log y = \frac{\log y}{\log \lambda^4n} \geq \frac{\log \lambda_60n}{\log \lambda^4n} = 15$$

(by the discussion following (2)) and observe first that from $1 - \rho_1 > 0$ follows

$$1 - \rho_1 = \frac{(1-\rho_1)\log y}{y} > 1 + (1-\rho_1)\log y = 1 + \Delta_1\chi_1^{-1}\log y$$

hence

$$\frac{\rho_1^{-1}}{y} < \frac{1}{1 + \phi\Delta_1} \quad (13)$$

To find an upper bound for $h_\chi(\rho_1^{-1})$ we write

$$h_\chi(\rho_1^{-1}) = h_\chi(0) + (\rho_1^{-1})h_\chi'(\xi) = 1 - (1-\rho_1)h_\chi'(\xi)$$

with some $\xi$ so $\rho_1^{-1} < \xi < 0$ (hence $\xi > -1/4$) and find a lower bound for $h_\chi'(\xi)$. But $h_\chi(s) = h_1(s)^2$ so $h_\chi'(\xi) = \xi h_{\chi^{-1}}'(\xi)h_1'(\xi) \geq \xi h_1'(\xi)$

and $h_1(s) = \frac{\sinh A^{-1}s}{A^{-1}s} = 1 + \frac{(A^{-1}s)^2}{3!} + \frac{(A^{-1}s)^4}{5!} + ...$ implies

$$h_1'(\xi) = \frac{1}{A}A^{-1}\xi + \frac{4(A^{-1}\xi)^3}{5!} + ... \geq \frac{1}{A}A^{-1}\xi$$

provided $A^{-1}\xi$ is small enough (noting $A^{-1}\xi < 0$ and that this appeal to the convergence
of the series can easily be made explicit), which by \(|\xi| < 1/4\), is the case when \(A \ll 1\) is given a suitable constant. Thus certainly
\(h_1'(\xi) > -\frac{1}{8A^2}\), from which the above inequalities show

\[h_{\xi}(\rho_{1-1}) < 1 + \frac{8}{8A^2}(1-\rho_1).\]  

(14)

Thus \(1 - h_{\xi}(\rho_{1-1})y^{\rho_{1-1}} > \frac{8A^2}{\phi \Delta_1} (1-\rho_1)^{\rho_1-1}\) by (13), (14) and since

\[\phi \Delta_1 - \frac{8}{8A^2}(1-\rho_1) = (1-\rho_1)(\log y - \frac{8}{8A^2}) \gg (1-\rho_1)\log y\]

by \(\frac{8}{8A^2} \ll 1\) and \(\log y \gg 1\) we have \(1 - h_{\xi}(\rho_{1-1})y^{\rho_{1-1}} \gg \frac{\phi \Delta_1}{1+\phi \Delta_1} \geq \frac{15\Delta_1}{1+15\Delta_1} \geq \frac{15\Delta_1}{1+30c_1}\), by \(\phi \geq 15\) and \(\Delta_1 \leq 2c_1\), and this proves

(12).

Combining (9), (10), (12) and noting that \(\Delta_1 \leq \Delta_1(H)\), then leaves us with

\[A_{H_1} \sum_{\rho \in \mathbb{C}} \frac{\log N\rho}{N\rho} \gg \Delta_1(H)(1-O(e^{-2c_1\log y})) - O\left(\frac{1}{(2n)!A^3}\right)\]

\[c_{22}n/A \leq \log ye^{-c_{22}nA} \ll Np < ye^{c_{22}nA}\]

and we can now construct \(c_{21}\): that is we choose \(c_{21}\) so large that
\[ c_{21} \geq (1-c_{17})^{-1}\left(\frac{8}{6} + 4\lambda + 7\right) \] and
\[ c_{21} \geq 8c_{18} \] and
\[ c_{21} \geq 2c_{1}^{-1}\log 2C \]

where \( C \) is the constant in the first \( O \)-term in (15). The first of these conditions ensures that (2) holds (since \( \xi \leq \left(\frac{8}{6} + 4\lambda + 7\right) \)), the second that (11) holds, and finally the third makes
\[ 1 - \theta(e^{-\frac{2c_{1}^{-1}\log y}{1}}) \geq 1/2 \] in (15).

Thus (15) is replaced by
\[ \frac{\log Np}{Np} = \Delta_{1}(H) - O\left(\frac{1}{(2n)!A^{3}}\right) \]

and we must apply theorem 1', in an essential way, for the first time. But first, if \( e_{1}(H) = 0 \) then \( \Delta_{1} \gg 1 \) so if \( A \gg 1 \) is given an appropriate value then
\[ \Delta_{1}(H) - O\left(\frac{1}{(2n)!A^{3}}\right) \gg 1 \]

which is sharper than anything we have claimed. Suppose \( e_{1}(H) = 1 \) hence \( \Delta_{1}(H) = \Delta_{1} \) (by the definition (15)) which, taking theorem 1'
in the form of its corollary, shows
\[ \Delta_{1}(H) \gg \min(c_{K}^{-1}, A^{-2}\log A^{4n}) \gg \min(c_{K}^{-1}, A^{-2}) \]
since $Q = A^{2n}$. Then if the minimum is attained at $c_K^{-1}$ the right side of (16) is

$$c_K^{-1}(1 - \frac{c_K}{(2n)!A^2}) \gg c_K^{-1}(1 - \frac{A^{-3}}{n^2}) \gg c_K^{-1}$$

for large enough $A \gg 1$, since $c_K \leq (2n)!$. And if the minimum is attained at $A^{-2}$ then the right side of (16) is

$$A^{-2}(1 - \frac{A^{-1}}{n^2}) \gg A^{-2}.$$

In either case we have proved lemma 13. Before going on to some better formulated versions we observe that the lower bound comes completely from the supposed exceptional zero, and that, if it did not exist then we would have $A^{-1}$ as lower bound: thus $A^{-1}$ can replace the right side unless there is an exceptional zero $(\rho_1, x_1)$ for $A^{2n}, A^2$ and even then we have $A^{-1}$ unless $x_1(H) = 1$ and $x_1(C) = 1$ (for in the proof of (12) we observed that there is "nothing to prove" i.e. the left side is $\gg 1$, unless $x_1(C) = 1$).

Also the conclusion of lemma 13 in the form of a minimum is needed in case $c_K = (2n)!$ actually occurs, for then $c_K^{-1}A^{-1}$ is the smaller term unless $A \geq n e^{O(n)}$, a condition which would make $y \geq n c_2 n^2 e^{O(n^2)}$ and is then unacceptable to us (in the sense that the condition $y \geq n e^{O(n)}$ is already considered as a flaw in the Introduction). Of course the corollary also shows that on the Artin conjecture (and in one special case) we can take $c_K = n$, eliminating the need for the minimum.

There is one important way in which it seems that it should be possible to improve the results: namely if $L/K$ is the class field
to \( H \) then we have

\[
\sum_{\chi(H)=1} \log d_\chi = \log d_L
\]

by the conductor-discriminant formula. In the derivation of (3) we did not try to take advantage of (17) and instead simply bounded \( d_\chi \) by \( A^{2n} \) for all \( \chi \) thus introducing the \( h_H \) terms. This leads us to speculate that lemma 13 should be true with the condition \( A^n \geq d(H)^{1/2} \) replaced by something like

\[
A^n \geq \left( \prod_{\chi(H)=1} d_\chi \right)^{1/h_H} = d_L^{1/[L:K]}
\]

i.e. replacing the maximum by the geometric mean. This would lead to much nicer conclusions; however we are unable to make any real use of (17) because of the means of proof of theorem 4: specifically, the difficulty arises from the use of the Selberg sieve and the subsequent need to estimate the distribution of integral ideals, and manifests itself in the transition from lemma 8 to its corollary when we had to estimate

\[
\sum_{\chi(H)=1} d_\chi^{1/2}
\]

(essentially).

Finally, we can now be very precise about why the explicit formula of §1 is needed instead of the "usual" one: the reason is that we can absorb the \( h_H \) coefficient in the error terms of (3) by a condition of the form \( A^n \geq h_H^{O(1)} \) so that the validity of the lemma is proved already for \( y \geq h_H^{O(1)} \), whereas by the "usual" explicit formula we would need a condition like \( y \geq h_H^{O(n)} \) to get a conclusion. If we
take $H = P_m$ for some $m$ then the bounds (9) of III§2 can be quite sharp and, in these cases, our result is then weakened considerably.

We turn briefly to (possibly) non-primitive congruence class groups to show that most of the work is already done by lemma 13 (so amplifying the "third reason" of III§2)

**COROLLARY.** Let $H$ mod $m$ be a congruence class group and, in addition to the hypotheses of lemma 13 (which depend only on the equivalence class of $H$), suppose that $A \geq (\log N(n))^{1/n}$ where $n$ is the product of the primes $p$ dividing $m$ which do not divide $f_H$. Then the conclusion lemma 13 holds, with constants $c_{21}^1, c_{22}^1$.

**PROOF:** Let $H_x$ be the primitive congruence class group mod $f_H$ which induces $H$ (in the sense of III§2): then the equation (3) holds for $H_x$ and moreover the right side of the equation depends only on the equivalence class $\bar{H} = \bar{H}_x$. And the left side differs from the corresponding expression for $H$ by at most

$$h_H \sum_{p \in \mathcal{C}} \frac{\log Np}{Np} \frac{\log Np}{Np} \leq A_H \sum_{np > y} -e/A \frac{\log Np}{Np}$$

$$\leq y^{-1/\varepsilon} A_H \sum_{p/m, p \not\mid f_H} \log Np \leq h_H y^{-1/\varepsilon} \log N(n)$$

where the first inequality is by (10) of III§1 and the last by $\varepsilon \leq A$. By our hypothesis this quantity is $\leq h_H y^{-1}AA^2 \leq h_H y^{-1}A^2$ which is small enough that (4) still holds (for example, because $A^2 \leq y^{1/30}$.
by the discussion following (2), we can already absorb this in the
first error term of (3)) with a slightly larger constant in the 0-term
and the rest then goes as before. The condition \( A \geq (\log N(n))^{1/n} \)
is clearly extremely weak in comparison to the other conditions in
lemma 13, unless \( m \) contains some large "irrelevant" prime ideals.

We formulate a simple consequence of lemma 13 which shows more
clearly about how strong it is, and which we state as

**THEOREM 5.** There exist effective positive constants \( c_{23}, c_{24}, c_{25}, c_{26} \),
independent of \( K \), with the following property: if \( C \) is a coset of
a primitive congruence class group \( H \) of \( K \), and we put

\[
B_H = \max(d(H)^{c_{23}}, h_H^{c_{23}}, n^{8c_{23}}, e^{c_{24}}),
\]

and define \( v_C(x) = \) the number of prime ideals \( p \in C \) so that

\[
x < Np < x + x^{1-\frac{c_{25}}{n}};
\]

then we have

\[
v_C(x) \gg \min(c_{26}^{-c_{26}}, c_{26}^{-5c_{26}}) \frac{1-\frac{c_{25}}{n}}{h_H \log x}
\]

for every \( x \geq B_H \).

In particular, there is a prime ideal \( p \in C \) for which \( N(p) \leq 2B_H \).

**PROOF:** WE apply lemma 13 with

\[
\lambda = \frac{1}{2}, \theta = .16
\]

thus the constants \( c_{21}, c_{22} \) are determined. Let \( x \geq B_H \) and put
\[
A = x^{1/c_{21}^n}, \quad y = xe^{c_{22}^n/A}.
\]

then, putting

\[
c_{23} = \frac{1}{2} c_{21},
\]

we have

\[
x \geq B_H \text{ implies } A \geq d(H)^{2n_H} e^{2n_H/n}, \quad e^{c_{24}/c_{21}},
\]

and \( y = xe^{c_{22}^n/A} > x = A^{c_{21}^n} \),

so, if \( c_{24} \) is large enough, then lemma 13 applies and gives

\[
\begin{align*}
h_H \frac{\log x}{x} \sum_{p \in C} 1 & \geq \min(c^{-1}A^{-1}, A^{-3}) \\ x < N_p < xe^{c_{22}^n/A}
\end{align*}
\]

since if \( c_{24} \) is large enough then

\[
N_p > x \text{ implies } \frac{\log N_p}{N_p} \leq \frac{\log x}{x}.
\]

Now \( e^{c_{22}^n/A} - 1 \leq 2 \cdot \frac{2c_{22}^n}{A} \) provided \( \frac{2c_{22}^n}{A} \) is small enough:

but, for large enough \( c_{24} \), we have \( 4c_{22} \leq A^{1/4} \) and \( n \leq A^{1/4} \)

by (19), so \( \frac{4c_{22}^n}{A} \leq \frac{1}{4} \frac{1}{4} = A^{-1/2} \) and in particular, \( \frac{2c_{22}^n}{A} \) is small if \( A \) is large (which is ensured by (19), for large enough \( c_{24} \)). But then
on putting \( c_{25} = \frac{1}{2c_{21}} \).

Finally (20) becomes

\[
v_C(x) \geq \min(c_{1-\frac{1}{n}}c_{25/n}, A^{-3}c_{25/n}) \frac{1-c_{25}/n}{h_H \log x}
\]

when, putting \( c_{26} = \frac{1}{2c_{21}} (= c_{25}) \), we have

\[
A^{-1}x^{c_{25/n}} = x^{2c_{25/n}} - c_{25/n} = x^{c_{26/n}} \quad \text{and}
\]

\[
A^{-3}x^{c_{25/n}} = x^{2c_{25/n}} - c_{25/n} = x^{5c_{26/n}} \quad ,
\]

as claimed. The "in particular" follows on taking \( x = 2^x \) and \( 1-c_{25/n} \)

noting that \( x + x \leq 2x \).

We have written the statement of the theorem in the form above because it seems that

\[
\lim_{x \to \infty} \frac{v_C(x)}{x^{1-c_{25}/n}} = 1
\]

\[
\frac{x}{h_H \log x}
\]

should hold; of course this (and much more) can be proved assuming the Generalized Riemann Hypothesis. Also we have kept \( c_{26} \) distinct from \( c_{25} \) because if we apply lemma 13 with \( \lambda = \frac{1}{2} \) and \( \theta > 16 \) then,
as $\theta$ becomes larger, $c_{26}$ will be smaller in comparison to $c_{25}$. However, since we have not kept explicit track of the dependence of our constants on the parameters $\lambda, \theta$ of theorem 4, it is no longer possible to see how $c_{25}$ behaves as $\theta$ becomes larger.

An interesting special case of theorem 5 is handled in

**COROLLARY.** Every ideal class of $K \neq \mathbb{Q}$ contains a prime ideal $p$ with

$$Np \leq \max(d_K^n, c_{27}^{28n})$$

**PROOF:** Apply theorem 5 to cosets $H = P_0$ (so actually considering ideal classes in the strict sense; however the division into broad ideal classes is coarser). Then $d(H) = d_K$, $h_H \leq e^{\mathcal{O}(n)} d_K$ (by 21) of §2, with $a = 1$) and Minkowski's lower bound $d_K \geq C^n$ with $C > 1$ allows us to absorb the $e^{\mathcal{O}(n)}$ and $e^{c_{24}^{2n}}$ into $d_K^{c_{27}}$, provided $K \neq \mathbb{Q}$. When $K = \mathbb{Q}$, $d_K = 1 = n$ makes the result false, although it is not interesting then anyway.

For the sake of completeness we point out that the (weaker) version of theorem 5 stated in the introduction follows from theorem 5 by using the inequalities (9) of III§2 for $d(H)$, $h_H$ (clearly $d(H) \leq D_H$) and then employing the Minkowski lower bound for discriminants to absorb all of the $e^{\mathcal{O}(n)}$ terms.
§3 Prime ideals in Tchebotarev classes

We apply theorem 5 in the setting of (possibly non-abelian) galois extensions of number fields using entirely algebraic methods, involving an essential use of class field theory (those uses up till now can be eliminated) and the theory of the Artin conductor, to formulate theorem 5 in terms of the irreducible representations of the galois group.

Let L/F be a finite galois extension of algebraic number fields with galois group G. We repeat a few definitions, for convenience:

if $P$ is a prime of L unramified over F then there is a unique automorphism $\sigma \in G$ so that $P^\sigma = P$ and so that

$$\alpha^\sigma \equiv \alpha^{N_p} \mod P$$

(1)

for every algebraic integer $\alpha$ of L, where $p$ is the prime of F lying below $P$ and where $N_p$ is the absolute norm = the cardinality of the residue class field at $p$. This $\sigma$ is the Frobenius automorphism at $P$, and is denoted $(P, L/F)$.

On the other hand, if $p$ is a prime of F unramified in L then the primes $P$ of L lying above $p$ are all conjugate (under $G$) so the Frobenius automorphisms $(P, L/F)$ fill a whole conjugacy class of $G$: we denote this conjugacy class by $(p, L/F)$.

Then we can state the problem we will consider: given a conjugacy class $C$ of $G$ we call the set of primes $p$ of F satisfying $(p, L/F) = C$ the Tchebotarev class associated to $C$, and recall that the Density Theorem of Tchebotarev [28] states that the (Dirichlet) density of the Tchebotarev class associated to $C$ exists and equals $|C|/|G|$
(where we write $|X|$ for the cardinality of the set $X$). Our problem then is to show the existence of a prime $p$ in this Tchebotarev class associated to $C$ with "small" absolute norm.

We proceed to do this by the reduction of Deuring [5] to abelian extensions (see p. 169 of Lang [18]); note that Serre has given an alternative approach to this argument using the orthogonality relations for characters (see Lagarias-Odlyzko [16]), however we only need part of Deuring's argument (which we therefore give).

Let $L/F$, $G$, and $C$ be as above, fix $\sigma_0 \in C$ and let $A$ be any abelian subgroup of $G$ containing $\sigma_0$. Let $K$ be the subfield of $L$ fixed by $A$, hence $A = \text{gal}(L/K)$. Now $L/K$ is abelian so $L$ is a class field over $K$ and there is a primitive congruence class group $H$ of $K$ so that the Artin map

$$I(q_H)/H \to \text{gal}(L/K) = A$$

induced by $q \to (q,L/K)$, for primes $q$ of $K$, is an isomorphism (note that, since $A$ is abelian, the conjugacy class $(q,L/K)$ reduces to a single element).

Suppose that we can find a bound $M$ (depending on $L/K$) and a prime $q$ of $K$ so that

$$(q,L/K) = \sigma_0, \ N_q \leq M, \ \deg q = 1, \ \text{and}$$

so if $p$ is the prime of $F$ lying under $q$ then $p$ is unramified in $L$. 
Here $\deg q = 1$ is as in the last paragraph of §1 (which will also allow us to prove the existence of $q$). Now if $q, M$ are found as in (3) and if $p$ is the prime of $F$ specified in (3) then

**Claim:** $(p, L/F) = C$ and $Np \leq M$.

By the last condition of (3), $p$ is unramified in $L$ so $(p, L/F)$ is defined. By $q/p$ the residue class field at $p$ is a subfield of the residue class field at $q$, while $\deg q = 1$ says the residue class field at $q$ is the prime field. Thus these residue class fields must coincide and we have $Nq = Np$ (since the absolute norm = cardinality of residue class field); in particular, $Np \leq M$. Now let $P$ be any prime of $L$ lying above $q$ (hence above $p$): then $(q, L/K) = \sigma_0$ says that $(P, L/K) = \sigma_0$ so $p^{\sigma_0} = P$ and $a^{\sigma_0} \equiv a^{Nq} \mod P$ for any integer $a$ of $L$. Since $Nq = Np$ this is precisely the statement (1) which means $(P, L/F) = \sigma_0$, where $\sigma_0 \in C$ implies $(p, L/F) = C$ and the claim is proved.

Thus we must return to the problem of finding $q$ as in (3) which we do by using the reciprocity law (2) and theorem 5; first, however, we need to make a slight adjustment to handle the last condition in (3).

Let $n$ be the product of all these primes $q'$ of $K$ which are unramified in $L$ but so that the prime $p'$ of $F$ lying under $q'$ is ramified in $L$ (since there are only finitely many such $p'$, and each has finitely many $q'$ lying over it, the ideal $n$ is well defined). Then letting $H'$ be the congruence class group $\mod (n_H')$ induced by $H$ and composing the canonical isomorphism
I(n\mathcal{H}_H)/H' \rightarrow I(\mathcal{H}_H)/H

(discussed in III§2) with the Artin isomorphism (2) gives the isomorphism

I(n\mathcal{H}_H)/H \xrightarrow{\omega} A

which is again called the Artin isomorphism (but with the interpretation modulus \(n\mathcal{H}_H\)). Because the prime factors of \(\mathcal{H}_H\) are precisely those primes \(q''\) of \(K\) which are ramified in \(L\) we see that \(n\) and \(\mathcal{H}_H\) are relatively prime.

We can apply theorem 5 with two modifications: first, by the last paragraph of §1, the condition \(\deg q = 1\) can be added to the conclusion of theorem 5 (after, possibly, increasing the constants), and second, we apply theorem 5 not to the primitive congruence class group \(H\) but rather to \(H'\). By the corollary to lemma 13 (with the above \(n\) equal to the \(n\) of the corollary, because "\(m'' = n\mathcal{H}_H\) with \(n\) prime to \(\mathcal{H}_H\)"

the only change necessitated by this in the statement of theorem 5 is that a term \((\log N(n))^{O(1)}\) must be included in the definition of "\(B_H\)" and that the constants may again need to be changed (because in the proof of theorem 5 the corollary to lemma 5 is applied to \(A = x^{1/c'_{21}n}\) and the condition \(A \geq (\log N(n))^{1/n}\) i.e. \(x \geq (\log N(n))^{c'_{21}}\) is then needed.

Thus we let \(C'\) be the coset mod \(H'\) so that \(w'(C') = \phi_o\) and we choose \(q \in C'\) and \(\deg q = 1\) by theorem 5 hence \(N(q) \leq M\) with

\[
M = 2 \max(d(H)^{O(1)}, h_H^{O(1)}, n_K^{O(n)}, e^{O(n)}), (\log N(n))^{O(1)})
\]
and the construction above shows immediately that this $q$ satisfies
(3) with $M$ given by (5), hence by the claim it remains only to
reinterpret $M$ in terms of the extension $L/F$. Note that $d(H), h_H$
being independent of the interpretation modulus (by III§2) is now
useful to us: this completes the "second reason" of III§2.

To begin we observe that by the Minkowski lower bound for discrim-
inants we have

$$e^n_K \leq d_K^0(1)$$

and since $d_K \leq d(H)$ we get

$$0(n_K^e) \leq d(H)^0(1)$$

(6)

so eliminating one of the terms in the maximum (5).

Moreover clearly

$$h_H = |A|$$

(7)

and

$$n_K^n = \frac{n_L}{|A|}$$

(8)

and since both of these term (7) (8) have the potential (in extreme cases)
of being larger than $d(H)^0(1)$ we accept them as they are.

Next we turn to $\log N(n)$: let $q'$ be a typical prime
factor of $n$ and let $p'$ be the prime of $F$ below $q'$. Since
$p'$ is ramified in $L$ and since $L/F$ is galois, every prime
factor $P'$ of $p'$ extended to $L$ must divide $\mathfrak{d}_{L/F}$. On the other hand every prime factor $P'$ of $q'$ extended to $L$ must not divide $\mathfrak{d}_{L/K}$ since $q'$ is unramified in $L$. Thus every prime factor $P$ of $q'$ extended to $L$ divides $\mathfrak{d}_{L/F}$ and not $\mathfrak{d}_{L/K}$ hence by:

$$\mathfrak{d}_{L/F} = \mathfrak{d}_{L/K} \mathfrak{d}_{K/F}$$  \hspace{1cm} (9)

we see that $q'$ divides $\mathfrak{d}_{K/F}$; so $n$, being a product of distinct $q'$, also divides $\mathfrak{d}_{K/F}$. Therefore

$$N(n) = N_{K/Q}(n) \leq N_{K/Q}(\mathfrak{d}_{K/F}) = N_{F/Q}(d_{K/F}) = d_{K/F}^{-[K:F]}$$

$$\leq d_K \leq d(H)$$

and

$$(\log N(n))^{(1)} \leq (\log d(H))^{(1)}$$  \hspace{1cm} (10)

Finally we come to the most important term $d(H)$ (because it is usually largest), and recall some facts about the Artin conductor, for which see Artin-Tate [1]; given a finite galois extension $L/F$ with galois group $G$ and the character $\theta$ of a representation of $G$ the Artin conductor $\mathfrak{d}(\theta)$ associated to $\theta$ is an integral ideal of $F$ which is defined formally. The Artin conductor has, however, suggestive functorial properties to which it owes its significance. For our purposes we need to know:
i) if $\theta, \theta'$ are characters of representations of $G$
then $\delta(\theta + \theta') = \delta(\theta)\delta(\theta')$

ii) if $A$ is a subgroup of $G$ with fixed field $K$ and
if $\chi$ is the character of a representation of $A$
then
$$\delta(\text{Ind}^G_A \chi) = d^{X(1)}_{K/F} N_{K/F} (\delta(\chi))$$

iii) if $A$ is abelian and if $\chi$ is a character of a
representation of $A$, define the congruence class
character $\psi$ by composing the Artin map 2) with
$\chi$. Then $\delta(\chi) = \delta_C$.

Returning to our original setting $F \subseteq K \subseteq L$ we see immediately
from iii) and 2) that

$$d(H) = \max \{d_K N(\delta_\chi): \chi \text{ congruence class character} \} \quad (11)$$
so $\chi(H) = 1 = \max \{d_K N_{K/Q} (\delta(\chi)): \chi \text{ character of an irreducible representation of } A\}.$

However $[K:F] = \frac{[L:F]}{|A|}$ so by ii) and the tower formula for
discriminants we have

$$d_F^{[A]} N(\delta(\text{Ind}^G_A \chi)) = d_{F/Q}^{[K:F]} N_{K/F} (d_K N_{K/F} (\delta(\chi)))$$

$$= d_F^{[K:F]} N_{F/Q} (d_{K/F} N_{K/Q} (\delta(\chi))) = d_N (\delta(\chi))$$
since a abelian implies $x(l) = 1$, i.e.

$$d \frac{[L:F]}{K} N(\pi(x)) = d \frac{|A|}{F} N(\pi(\text{Ind}_A^G x)).$$

(12)

But putting together 5), 6), 7), 8), 10), 11), and 12) we see that we have proved

**THEOREM 6.** Let $L/F$ be a finite galois extension of algebraic number fields with galois group $G$, and let $C$ be a conjugacy class of $G$. For an abelian subgroup $A$ of $G$ define

$$B_{L/F}(A) = \max \{ N(\pi(\text{Ind}_A^G x)) : x \text{ is an irreducible character of } A \}$$

If $A \cap C$ is not empty then there exists a prime $p$ of $F$ in the Tchebotarev class associated to $C$ satisfying

$$Np < 2 \max \left\{ \left( d \frac{|A|}{L/F}(A) \right)^{c_{29}}, \frac{n_L}{|A|} \right\}$$

**PROOF.** Remains only to point that if $L = F = Q$ then we need the factor 2.

The point of theorem 6 is that everything can be interpreted in terms of the extension $L/F$ and of the variable abelian subgroup $A$ (provided it meets $C$). This can be made more precise: for each character $x$ of $A$, $\text{Ind}_A^G x$ is the character of a representation of $G$ so
\[
\text{Ind}_A^G \chi = \sum \langle \text{Ind}_A^G \chi, \theta \rangle \theta = \sum \langle \chi, \text{Res}_G^A \theta \rangle \theta
\]

where \( \theta \) runs through the irreducible characters of \( G \) and

where \( \langle \chi, \text{Res}_G^A \theta \rangle \) is a non-negative integer. Thus by i) we

have that

\[
\delta(\text{Ind}_A^G \chi) = \prod_{\theta} \langle \chi, \text{Res}_G^A \theta \rangle
\]

depends on the Artin conductors of the irreducible representations

of \( G \) and on the non-negative integers \( \langle \chi, \text{Res}_G^A \theta \rangle \) which depend

on how the abelian subgroup \( A \) of \( G \) (meeting \( C \)) is situated

in \( G \). At any rate the size and location of \( A \) determines the

strength of theorem 6 to a large extent.

As an indication of how large \( d \left[ \frac{[L:F]}{|A|} \right] B_{L/F} (A) \) is we note

that from \( \sum \text{Ind}_A^G \chi = \text{Ind}_A^G 1 \) follows

\[
\prod_{\chi} \left(\frac{[L:F]}{|A|} \right) N(\delta(\text{Ind}_A^G \chi)) = d_L
\]

(13)

since the right side is then \( d \left[ \frac{[L:F]}{|A|} \right] N(\delta(\text{Ind}_A^G 1)) \) which equals

\( d \left[ \frac{[L:F]}{L/F} \right] N(d_{L/F}) = d_L \) by ii) and the usual discriminant formula.

From 13) we have

\[
d_{L/F} \left[ \frac{[L:F]}{|A|} \right] B_{L/F} (A) \geq d_L^{1/|A|}
\]

(14)
and can expect near equality if 

\[ \frac{|L:F|}{d_F|A|} N(\delta(\text{Ind}_A^G)) \]  

is almost constant as \( \psi \) varies. As an example where we come somewhat close to this inequality we note

\[ \frac{|L:F|}{d_F|A|} B_{L/F}(A) \leq \frac{1/\phi(|A|)}{d_F}, \]

if \( A \) is cyclic where \( \phi \) is Euler's \( \phi \)-function. This is immediate from 13) on observing \( \delta(\text{Ind}_A^G\psi) = d_{K/F}N_{K/F}(\delta(\psi)) \) takes the same value whenever \( \psi \) is a faithful character of \( A \) (this is easiest to see on changing from \( \delta(\psi) \) to \( \delta_x \) via iii) and noting that \( \delta_x \) depends only on \( \ker X \) and that there are \( \phi(|A|) \) faithful characters of \( A \).

§4 EXAMPLES

We content ourselves with two very simple examples. First we recall that the Minkowski theorem quoted in 21) of §2 asserts the existence of an integral ideal \( \alpha \) in each ideal class of \( K \) with \( N\alpha \leq c^n d_K^{1/2} \). From the proof of this result we are led to expect that this is reasonably sharp, at least in general. But the corollary to theorem 5 is a result of comparable quality (apart from the value of \( c_{27} \)) provided \( d_K \) is at least as large as \( n^0(n) \). In fortunate circumstances we can prove that this result is also reasonably sharp as in the

**EXAMPLE:** \( K = \mathbb{Q}(\zeta_\ell) \) the field of \( \ell \)-th roots of unity where \( \ell \) is a large prime. Then \( d_K = \ell^{\ell-2} \), \( n = \ell-1 \) so by the corollary
to theorem 5 every ideal class of $K$ contains a prime ideal $p$ with

$$N(p) \leq \ell^{c_{32}\ell}.$$  

Conversely suppose $X$ is the smallest number so that every ideal class contains a prime ideal with norm $\leq X$. Then

$$h_K \leq \sum_{Np \leq X} \sum_{Na \leq X} 1 \leq e^{O(n)}X^{3/2} = e^{O(\ell)}X^{3/2}$$

by 20) of §2 with $a = 2$, hence

$$X \geq e^{O(\ell)}h_K^{2/3}.$$  

Now from theorem 2 of Stark [27] it is easy to deduce the existence of $C > 0$ so that $h_K \geq \ell^{c_{32}}$ for sufficiently large $\ell$; thus

$$X \geq e^{O(\ell)}\ell^{2/3c_{32}} \geq \ell^{C'\ell}$$

for large enough $\ell$. In particular apart from the value of $c_{32}$ the application of theorem 5 gives a best possible result.

Finally it seems necessary to give an example for theorem 6 so we give the simplest possible; and already this will make apparent the limitations imposed by reduction to abelian extensions.

**EXAMPLE:** $L/Q$ a galois extension whose galois group $G$ is dihedral of order $2m$ where $m$ is odd. Let $A$ be the normal
abelian subgroup of $G$ of index 2: then $\text{Ind}_A^G \chi = 1 + \phi$ where $1, \phi$ are all the 1-dimensional irreducible characters of $G$, and if $\psi \neq 1$ is an irreducible character of $A$ then $\text{Ind}_A^G \psi$ is also an irreducible character of $G$ of degree 2. Since all irreducible characters of $G$ arise in this way we have

$$\text{B}_{L/Q}(A) = \max \{N(\delta(\theta)) : \theta \text{ an irreducible character of } G\}$$

so, by theorem 6, for every conjugacy class $C$ of $G$ which meets $A$ there is a prime $p$ of $\mathbb{Q}$ with $\left(p, L/Q\right) = C$ and $p \ll \max \left\{ \text{B}_{L/Q}(A)^{\epsilon_3}, m^{\epsilon_4} \right\}$

which seems a reasonably strong result.

On the other hand all of the elements of $G$ which lie outside $A$ form a single conjugacy class $C_\times$; all of the elements of $C_\times$ have order 2 and the only abelian subgroups of $G$ which meet $C_\times$ are cyclic of order 2. It is clear that the result of theorem 6 is very weak in this situation, especially when $m$ is large. In particular the bound we get is much weaker than the corresponding bound for conjugacy classes which meet $A$.

To make this state of affairs more dramatic observe that the Tchebotarev class associated to $C_\times$ has Dirichlet density $1/2$ whereas for classes $C$ which meet $A$ the corresponding Dirichlet densities are $\frac{1}{2m}$ (for the unit class) and $\frac{1}{m}$ (for each of the $\frac{m-1}{2}$ others).
LIST OF REFERENCES

1. Artin, E. and Tate, J., Class Field Theory, Benjamin, New York (1968).
4. Davenport, H., Multiplicative Number Theory, Markham, Chicago (1967).


