Project Report

An Introduction to Elementary Methods in Analytic Number Theory

Summer, 2023

Abhisruta Maity[1]
BS-MS Student, Year 2
Indian Institute of Science Education and Research, Kolkata

Under the Guidance of:

Dr. Satadal Ganguly[2]
Associate Professor, *Indian Statistical Institute, Kolkata, India*

[1] Email ID: am21ms006@iiserkol.ac.in

[2] Email ID: sgisical@gmail.com

Acknowledgement

I extend my gratitude to first and foremost, my supervisor Dr. Satadal Ganguly for his crucial guidance. The meetings with him were really constructive and inspiring. My perspective towards problem-solving changed a lot through the interactions with him.

I am grateful to my parents for their constant support and their curiosity towards mathematics. I am also grateful to my friends who have taught me how to go beyond the barrier of oneself.

Abstract

This project summarizes two introductory branches of analytic number theory. In the first part, we have studied multiplicative functions and their averages. Dirichlet Hyperbola method brings better estimates; So selected applications were studied. We have attempted to understand the equivalence of Prime Number Theorem with mean values of Mobius function. Selberg's asymptotic formula and related statistics were studied. Part two of the project corresponds to the understanding of the Dirichlet characters through some methods from Fourier analysis on finite groups. A short introduction to Gauss sums is provided. Finally, we conclude with Polya-Vinogradov inequality.

Contents

I

Elementary Estimates

1 Dirichlet Convolution and Asymptotics

Dirichlet convolution is a binary operation between two arithmetic functions to produce another arithmetic function. This operation naturally arises in number theory. Before we understand it, we first define some identity functions.

Definition 1.0.1 (Identity Functions)

Let n be a positive integer. We define "Kronecker delta" on $\mathbb{N}$:

$$\delta(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise} \end{cases} = \left\lfloor \frac{1}{n} \right\rfloor.$$

Next, we define characteristic function on $\mathbb{N}$, denoted by $1_{\mathbb{N}} = 1 = u$:

$$1(n) = 1.$$

And last, we define identity map on $\mathbb{N}$, denoted by $\text{id} = N$:

$$\text{id}(n) = n.$$

In analytic number theory, one of the important problems is to find "distibutions" of certain subsets in the set of natural numbers. The functions discussed above are the characteristic functions of respective subsets. These functions keep appearing while evaluating partial sums of arithmetic functions.

§1.1 Dirichlet Convolution

We begin with the definition:

Definition 1.1.1 (Dirichlet Convolution)

Let $\mathcal{F}$ denote the family of arithemtic functions, i.e., complex-valued functions defined on $\mathbb{N}$. Define $* : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ such that for every positive integer n,

$$f * g(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right).$$

Clearly, $f * g$ is an arithmetic function produced via f and g, which is formally called the *Dirichlet convolution* or *Dirichlet product* of f and g. The name also refers to the binary operation $*$.

Let us state a few properties of $*$ in $\mathcal{F}$.

1. $\mathcal{F}$ is closed under $*$.
2. $*$ is associative.
3. $*$ is commutative.
4. Kronecker delta δ on $\mathbb{N}$, as defined in Definition 1.1.1, acts as the *identity* function under the operation $*$ in $\mathcal{F}$.

The above properties of $*$ can be verified easily by exploiting the definition of it. Perhaps at this point, it's very natural question to ask whether inverse of an arithmetic function exists under the same operation. The answer is *yes*, but only under a condition:

Theorem 1.1.2

Let f be an arithmetic function with $f(1) \neq 0$. Then there exists a unique arithmetic funtion g (called the *Dirichlet inverse* of g) satisfying:

$$f * g = \delta.$$

Moreover, g can be determined recursively:

$$g(1) = \frac{1}{f(1)}, \quad g(n) = -\frac{1}{f(1)} \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) g(d).$$

Proof. For a given arithmetic function f, define an arithmetic function g recursively as illustrated in the statement above. We will induct on n to show that for every positive integer n:

$$f * g(n) = \delta(n).$$

For $n = 1$, $f * g(1) = f(1)g(1) = f(1) \cdot \frac{1}{f(1)} = 1 = \delta(1)$. Assume that for some $k > 1$, the hypothesis holds for all $n < k$. Then

$$f * g(k) = f(1)g(k) + \sum_{\substack{d|k \\ d<k}} f\left(\frac{k}{d}\right) g(d) = f(1)g(k) - f(1)g(k) = 0 = \delta(k),$$

by induction hypothesis. Uniqueness can be justified by observing that the equation$f * g(k) = \delta(k)$ has a unique solution for $g(k)$, if $g(n)$ for each $n < k$ is known. Moreover, the solution is exactly what is show in the assertion. □

Hence, the family

$$\mathcal{F}^* = \{f \in \mathcal{F} : f(1) \neq 0\}$$

indeed forms a group under Dirichlet convolution $*$. Although, the group is not very interesting to study, still that structure gives us to use some element-wise properties (e.g. we can take $(f * g)^{-1} = g^{-1} * f^{-1} = f^{-1} * g^{-1}$, for granted.)

The central theme of this chapter is the following questions about an arithmetical function f:

- Is f multiplicative (we shall define the word in the next section)? If not, whether its “interaction” is good with other multiplicative functions.
- What is $f * 1$? In general, what is $f * g$?
- What is the asymptotic behaviour of certain “weighted" partial sums of f? Is there any exact formula?
- Is there any relation between different weighted sums?

Theorem 1.1.3 ($1_{\mathbb{N}}^{-1} = \mu$)

We have the following identity:

$$\mu * 1 = \delta = 1 * \mu$$

i.e., for all positive integers n,

$$\sum_{d|n} \mu(d) = \delta(n).$$

Here, μ denotes the Mobius function.

Proof. For $n = 1$, the equality holds trivially. For $n > 1$, the equality follows from binomal expansion of $(1-1)^k$, where k is the number of distinct prime factors of n. □

The preceding theorem allows us to evaluate a sum restricted to co-prime integers with respect to some modulus.

Example 1.1.4 (Exercise 3.12, [Apo98])

For real $s > 0$ and $k \geq 1$, we have

$$\sum_{\substack{n\leq x\\(n,k)=1}} \frac{1}{n^s} = \sum_{n\leq x} \frac{1}{n^s}\left[\frac{1}{(n,k)}\right] = \sum_{n\leq x}\frac{1}{n^s}\sum_{d|(n,k)}\mu(d) = \sum_{n\leq x}\frac{1}{n^s}\sum_{\substack{d|k\\d|n}}\mu(d)$$

To investigate the asymptotic formula, we first rearrange the finite sum. Each term of the sum involves $\mu(d)$ and n, where d is a divisor of k, and $n \leq x$ is a multiple of d. To rewrite the sum, we first fix a divisor d of k. All the multiples $n = dq$ such that $d \leq n \leq x$ if and only if the quotients $1 \leq q \leq x/d$. Hence we sum over all such divisors d and all such quotients q.

$$\sum_{n\leq x}\frac{1}{n^s}\sum_{\substack{d|k\\d|n}}\mu(d) = \sum_{d|k}\sum_{q\leq x/d}\frac{\mu(d)}{(dq)^s} = \sum_{d|k}\frac{\mu(d)}{d^s}\sum_{q\leq x/d}\frac{1}{q^s}.$$

which can be estimated asymptotically using the results due to Euler summation formula. For $s = 1$:

$$\sum_{d|k}\frac{\mu(d)}{d}\sum_{q\leq x/d}\frac{1}{q} = \sum_{d|k}\frac{\mu(d)}{d}\log\left(\frac{x}{d}\right) + \gamma\sum_{d|k}\frac{\mu(d)}{d} + O\left(\sum_{d|k}\frac{1}{d}\frac{d}{x}\right),$$

therefore the final expression looks like

$$\log x\frac{\varphi(k)}{k} + C_{k,1} + O\left(\frac{1}{x}\right).$$

For $s \neq 1$, we reach at

$$\frac{x^{1-s}}{1-s}\frac{\varphi(k)}{k} + C_{k,s} + O\left(\frac{1}{x^s}\right),$$

where $C_{k,s}$ is a constant depending upon k, s. Next we will assert a very useful theorem:

Theorem 1.1.5 (Mobius Inversion)

Let f and g be two arithmetical functions. We will have

$$f = g * 1,$$

if and only if,

$$g = \mu * f.$$

Proof. The proof is immediate by using Theorem 1.1.3. □

§1.2 Multiplicative Functions

Studying the multiplicative functions on integers is the heart of analytic number theory.

Definition 1.2.1 (Multiplicative Functions)

An arithmetical function f, which is not identically zero, is said to be *multiplicative* if for every $m, n \in \mathbb{N}$ with $(m, n) = 1$, we have

$$f(mn) = f(m)f(n).$$

A multiplicative function f is said to be *completely multiplicative* if for every $m, n \in \mathbb{N}$, we have

$$f(mn) = f(m)f(n).$$

Let's note a few points:

- Observe that, since multiplicative f is not identically zero, choosing c, such that $f(c) \neq 0$, we must have $f(c) = f(c \cdot 1) = f(c)f(1)$, which leads to $f(1) = 1$.
- The function $\mathrm{id}_\alpha(n) = n^\alpha$, is completely multiplicative for every complex number α. Note that, $\mathrm{id}_0 = 1$, the characteristic map on $\mathbb{N}$, and $\mathrm{id}_1 = \mathrm{id}$, the identity map on $\mathbb{N}$.
- Kronecker delta on $\mathbb{N}$ is also completely multiplicative.
- Mobius function μ is multiplicative, but not completely multiplicative.
- Liouville's function λ is completely multiplicative.
- If f, g are multiplicative (resp., completely multiplicative), then so is their point-wise product fg and quotient f/g, of course, only when $g(n) \neq 0$ for all n.

Theorem 1.2.2

Given an arithmetic function f with $f(1) = 1$. Then

1. f is multiplicative, if and only if,

$$f\left(\prod_i p_i^{\alpha_i}\right) = \prod_i f\left(p_i^{\alpha_i}\right),$$

for any set of primes p_i and any set of positive exponents α_i.

2. Multiplicative f is completely multiplicative, if and only if,

$$f(p^\alpha) = f(p)^\alpha,$$

for any prime p and any positive exponent α.

Proof. The proofs are just expanding and contracting the definitions. □

Theorem 1.2.3

If f and g are multiplicative functions, then $f * g$ is so.

Proof. The key idea: if $d|mn$, where $(m, n) = 1$, we can then write $d = ab$ uniquely, where $a|m$ and $b|n$. □

As a consequence, we have the following:

- Euler's Totient function, $\varphi = \mu * \mathrm{id}$, is multiplicative. It's not completely multiplicative (check $\varphi(4)$).

- The function $f(n) = \lfloor\sqrt{n}\rfloor - \lfloor\sqrt{n-1}\rfloor$, is multiplicative! One can easily check that, it's a square indicator function and thus it's multiplicative. One could also notice that $f = \lambda * 1$, hence could infer its multiplicativity. This idea can be applied in general setting.

Theorem 1.2.4

If f is multiplicative and $f * g$ is multiplicative, then g is multiplicative.

Proof. We will prove the contrapositive version of the statement. Suppose, f is a multiplicative function. Assume that, g is another arithmetic function which is not multiplicative. That means, there exists pair of positive integers m, n with $(m, n) = 1$, such that

$$g(mn) \neq g(m)g(n).$$

Choose *such* a pair m, n so that their product mn is smallest. This would mean that for any pair of positive integers $a < m$ or $b < n$, we would have $ab < mn$. Thus, for such pairs

$$g(ab) = g(a)g(b).$$

Now we are going to prove that $f * g$ is not multiplicative. We pick the above integers m, n and notice that,

$$f * g(mn) = \sum_{d|mn} f(d)g(mn/d) = \sum_{\substack{d|mn \\ d>1}} f(d)g(mn/d) + f(1)g(mn).$$

Since, f is multiplicative, $f(1) = 1$. Secondly, since $(m, n) = 1$, for $d|mn$ we must have, $d = ab$ with $a|m$ and $b|n$. Furthermore, $d > 1$ forces $a > 1$ or $b > 1$. With $(\frac{m}{a}, \frac{n}{b}) = 1$ and $\frac{m}{a} \cdot \frac{n}{b} < mn$, we get:

$$\sum_{\substack{d|mn \\ d>1}} f(d)g(mn/d) = \sum_{\substack{ab|mn \\ ab>1}} f(ab)g(mn/ab) = \sum_{a|m} f(a)g(m/a) \sum_{b|n} f(b)g(n/b) - g(m)g(n).$$

Therefore,

$$f * g(mn) = (f * g)(m) \cdot (f * g)(n) + g(mn) - g(m)g(n)$$

Since, $g(mn) - g(m)g(n) \neq 0$, the convolution $f * g$ is not multiplicative. □

Corollary 1.2.5

If f is multiplicative, so is its Dirichlet inverse f^{-1}.

Proof. Since, f is multiplicative and $f * f^{-1} = \delta$ is multiplicative, by the previous theorem, f^{-1} must be multiplicative. □

Theorem 1.2.6

Let f be completely multiplicative. Suppose, g and h are two multiplicative functions. Then, the point-wise product is distributive over Dirichlet product.

$$f(g * h) = fg * fh.$$

Proof. For any positive integer n,

$$f(n)(g * h)(n) = f(n) \sum_{d|n} g(d)h(n/d) = \sum_{d|n} f(d)g(d)f(n/d)g(n/d) = (fg * fh)(n),$$

since f is completely multiplicative. □

Theorem 1.2.7

Let f be multiplicative. Then f is completely multiplicative, if and only if,

$$f^{-1} = \mu f.$$

In particular,

$$1^{-1} = \mu.$$

Proof. Forward implication is a direct corollary of Theorem 1.2.6. We shall prove the other direction. Assume that,

$$f^{-1}(n) = \mu(n) f(n).$$

Then for any prime p and exponent $\alpha \geq 1$, we have

$$0 = \delta(p^\alpha) = \mu f * f(p^\alpha) = \sum_{d \mid p^\alpha} \mu(d) f(d) f(p^\alpha/d) = f(p^\alpha) - f(p) f(p^{\alpha-1}),$$

which implies,

$$f(p^\alpha) = f(p)^\alpha$$

by induction, proving that multiplicative function f is completely multiplicative. □

Example 1.2.8

Application of above theorems can be sketched like the following:

- $\varphi^{-1} = (\mu * \mathrm{id})^{-1} = \mu^{-1} * \mathrm{id}^{-1} = 1 * \mathrm{id}\mu$. In other words, $\varphi^{-1}(n) = \sum_{d|n} d\mu(d)$.
- $\sigma_\alpha^{-1} = (id_\alpha * 1)^{-1} = \mu * \mathrm{id}_\alpha \mu$. In other words, $\sigma_\alpha^{-1}(n) = \sum_{d|n} d^\alpha \mu(d) \mu(n/d)$.
- Dirichlet inverse of the square-indicator function $f^{-1} = (\lambda * 1)^{-1} = \mu * \mu\lambda = \mu * \mu^2 = \mu * |\mu|$.

§1.3 Summation Techniques

Our goal in this section is to brief the standard techniques used to convert sums into integrals. The proofs of the following theorems can be found in [Apo98].

Theorem 1.3.1 (Euler's Summation Formula)

If f has a continuous derivative f' on the interval $[y, x]$, where $0 < y < x$, then

$$\sum_{y < n \leq x} f(n) = \int_y^x f(t)\, dt + \int_y^x \{t\} f'(t)\, dt + \{x\} f(x) - \{y\} f(y).$$

Definition 1.3.2 (Riemann Zeta Function; Positive Real Argument)

Define $\zeta : \mathbb{R}_+ \setminus \{1\} \to \mathbb{C}$ in the following way:

$$\zeta(s) = \begin{cases} \sum_{n=1}^\infty \frac{1}{n^s} & \text{if } s > 1 \\ \lim_{x \to \infty} \left(\sum_{n \leq x} \frac{1}{n^s} - \frac{x^{1-s}}{1-s} \right) & \text{if } 0 < s < 1. \end{cases}$$

From Euler summation formula, one can work out the following asymptotic formulas:

1. $\sum_{n \leq x} \frac{1}{n} = \log x + \gamma + O(\frac{1}{x})$.

2. $\sum_{n\le x}\frac{1}{n^s}=\frac{x^{1-s}}{1-s}+\zeta(s)+O(\frac{1}{x^s})$, for $s>0, s\neq 1$.
3. $\sum_{n>x}\frac{1}{n^s}=O(x^{1-s})$, for $s>1$.
4. $\sum_{n\le x}n^\alpha=\frac{x^{\alpha+1}}{\alpha+1}+O(x^\alpha)$, for $\alpha\ge 0$.

Theorem 1.3.3 (Abel's Summation Lemma)

§1.4 How Much of a Forest Can Be Seen?

In analytic geometry, lattice points in x-y plane are those points whose coordinates are integers. An interesting class of problems arises with the lattice points which are "visible" from origin. But what exactly is this visibility?

Definition 1.4.1 (Visibility of Lattice Points)

Two lattice points P and Q are said to be mutually *visible* if the line segment joining them does not contain any lattice point other than P and Q.

Theorem 1.4.2

Two lattice points (x_1, y_1) and (x_2, y_2) are mutually visible, if and only if, $x_2 - x_1$ and $y_2 - y_1$ are relatively prime.

Remark 1.4.3. To distinguish, a vector (x, y) from (x, y) being the greatest common divisor of two integers x and y, we will use $\gcd(x, y)$ explicitly for the latter.

Proof. Notice that, (x_1, y_1) and (x_2, y_2) are mutually visible, if and only if, $(x_2 - x_1, y_2 - y_1)$ is visible from origin (i.e., $(x_2 - x_1, y_2 - y_1)$ is visible from origin; Of course, visibility relation is reflexive, symmetric but not transitive!). Therefore, it is sufficient to prove the assertion for (m, n) and origin $(0, 0)$.

Suppose that, (m, n) is visible from origin. Let $d = \gcd(m, n)$. We want to show that $d = 1$. If $d > 1$, then the lattice point shown as a vector,

$$\begin{bmatrix} m/d \\ n/d \end{bmatrix} = \frac{1}{d}\begin{bmatrix} m \\ n \end{bmatrix},$$

is lying on the line segment joining origin and (m, n). But that's a contradiction to our hypothesis that (m, n) is visible.

For the converse implication, assume that, $\gcd(m, n) = 1$. Suppose, (m', n') is a lattice point falling in the line segment joining origin and (m, n). Then $m' = \lambda m$ and $n' = \lambda n$ for some $0 < \lambda < 1$. First observe that, $\lambda \in \mathbb{Q}$, so let $\lambda = \frac{r}{s}$, with $\gcd(r, s) = 1$. Then,

$$m's = mr, \qquad n's = nr,$$

implies, $s|m$ and $s|n$ by Euclid's lemma. But this forces, $s = 1$, otherwise we would have $\gcd(m, n) \ge s > 1$, which is again a contradiction. □

Wait, how all these are relevant to the name of this section? Well, consider the x-y grid. Place a tree at each of the lattice points. You, as an observer, sit on the origin. By the theorems we saw earlier, you might have already guessed that, nature won't allow all the trees to be seen by you (unless, you move to the other dimension!). Now the question is, how much of the complete forest can the observer see? The answer is, indeed with much of a suprise, very simple but elegant: $\frac{6}{\pi^2}$! To prove this, we prove the following lemma:

Lemma 1.4.4 (Partial Sum of Euler's Totient Function)

For $x \geq 1$, we have

$$\sum_{n \leq x} \varphi(n) = \frac{3}{\pi^2} x^2 + O(x \log x).$$

Proof. For $x \geq 1$,

$$\begin{aligned}
\sum_{n \leq x} \varphi(n) &= \sum_{n \leq x} (\mu * \mathrm{id})(n) \\
&= \sum_{n \leq x} \sum_{d \mid n} \mu(d) \frac{n}{d} \\
&= \sum_{d \leq x} \mu(d) \sum_{q \leq x/d} q \\
&= \sum_{d \leq x} \mu(d) \left\{ \frac{x^2}{d^2} + O\left(\frac{x}{d}\right) \right\} \\
&= \frac{x^2}{2} \sum_{d \leq x} \frac{\mu(d)}{d^2} + O\left(x \sum_{d \leq x} \frac{1}{d} \right) \\
&= \frac{x^2}{2} \sum_{d \leq x} \frac{\mu(d)}{d^2} + O(x \log x).
\end{aligned}$$

Note that, $\lim_{x \to \infty} \sum_{d \leq x} \frac{\mu(d)}{d^2}$ exists. In fact, it is equal to $\frac{1}{\zeta(2)} = \frac{6}{\pi^2}$. Therefore,

$$\begin{aligned}
\sum_{n \leq x} \varphi(n) &= \frac{x^2}{2} \sum_{d \leq x} \frac{\mu(d)}{d^2} + O(x \log x) \\
&= \frac{x^2}{2} \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} - x^2 \sum_{d > x} \frac{\mu(d)}{d^2} + O(x \log x) \\
&= \frac{3}{\pi^2} x^2 + O(x \log x).
\end{aligned}$$

□

Remark 1.4.5. The bound can be improved by being careful in an intermediate step. However, sharper bound is not required for our purpose.

Consider a square region, centered at origin such that

$$|x| \leq r, \quad |y| \leq r.$$

Let $N(r)$ denote the number of lattice points contained by the square region. Suppose, $V(r)$ denotes the number of lattice points in that same square, which are visible from origin.

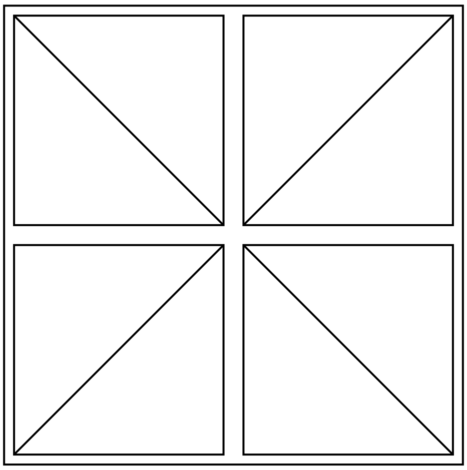
Theorem 1.4.6 (Density of Forest Visible from Origin)

We have,

$$\lim_{r \to \infty} \frac{V(r)}{N(r)} = \frac{6}{\pi^2}.$$

In other words, around 60.8% of the forest is visible from origin.

Proof. The nearest 8 lattice points are clearly visible from origin. Then observe that, a square can be broken into 4 squares, each of which can further be broken into 2 triangular pieces (shown in the figure below).



Therefore, to count $V(r)$ we make use of the following triangle:

$$\{(x,y) : 2 \leq x \leq r, \quad 1 \leq y \leq x\}.$$

We thus obtain:

$$V(r) = 8 + 8 \sum_{2\leq x\leq r} \sum_{\substack{1\leq y\leq x \\ (x,y)=1}} 1 = 8\sum_{x\leq r} \varphi(x) = \frac{24}{\pi^2}r^2 + O(r\log r),$$

where we have used Lemma 3.1.1. In contrast,

$$N(r) = (2\lfloor r \rfloor + 1)^2 = 4r^2 + O(r).$$

As a consequence, for some $K_1, K_2 > 0$,

$$\frac{\frac{6}{\pi^2} - \frac{K_1 \log r}{r}}{1 + \frac{K_2}{r}} \leq \frac{V(r)}{N(r)} \leq \frac{\frac{6}{\pi^2} + \frac{K_1 \log r}{r}}{1 - \frac{K_2}{r}},$$

for any $r \geq 1$. By squeeze theorem therefore,

$$\lim_{r\to\infty} \frac{V(r)}{N(r)} = \frac{6}{\pi^2}.$$

□

We state another beautiful result related to this setup.

Theorem 1.4.7

The set of lattice points in the x-y plane visible from origin contains arbitrarily large square gaps. Formally speaking, given an integer $q > 0$, there exists a lattice point (a,b) such that all of the lattice points

$$(a+r, b+s), \quad \text{with } 1 \leq r \leq q, 1 \leq s \leq q$$

are not visible from origin.

Proof. Consider a $q \times q$ array, whose entries are all distinct primes p_{ij}, $1 \leq i,j \leq q$. Let

$$m_i = \prod_{j=1}^{q} p_{ij} \text{ (row product)}, \quad M_j = \prod_{i=1}^{q} p_{ij} \text{ (column product)}$$

and let $M = \prod_{i=1}^q m_i = \prod_{j=1}^q M_j$. Observe that, m_is are pairwise relatively prime. So are M_js. Hence, the following simultaneous congruences:

$$x \equiv -i \bmod m_i,$$

for $1 \le i \le q$ and

$$y \equiv -j \bmod M_j,$$

for $1 \le j \le q$, will have solution a for (x, y), guaranteed by Chinese Remainder theorem. Now if we consider the lattice point $(x+i, y+j)$ with $1 \le i, j \le q$, then it will not be visible from origin. The reason is, $m_i | x+i$ and $M_j | y+j$ via construction. But there is exactly one prime p_{ij} which divides both m_i and M_j. This prime would force $\gcd(x+i, y+j) > 1$. □

§1.5 Dirichlet Hyperbola Method

Theorem 1.5.1 (Boundary Cases of Dirichlet Hyperbola Method)
Let f, g be Arithmetic Functions. Suppose $h = f * g$ and

$$F(x) = \sum_{n \le x} f(n); \quad G(x) = \sum_{n \le x} g(n); \quad H(x) = \sum_{n \le x} h(n)$$

Then

$$H(x) = \sum_{n \le x} f(n) G\left(\frac{x}{n}\right) = \sum_{n \le x} g(n) F\left(\frac{x}{n}\right)$$

Proof. We have,

$$\begin{aligned} H(x) = \sum_{n \le x} h(n) &= \sum_{n \le x} \sum_{d | n} f(d) g\left(\frac{n}{d}\right) \\ &= \sum_{n \le x} \sum_{q, d : qd = n} f(d) g(q) \\ &= \sum_{d \le x} \sum_{q \le x/d} f(d) g(q) \\ &= \sum_{d \le x} f(d) \sum_{q \le x/d} g(q) = \sum_{d \le x} f(d) G\left(\frac{x}{d}\right) \end{aligned}$$

The other part is almost identical to it. □

In particular, we have the following results:

- For $g = 1$, we have

$$H(x) = \sum_{n \le x} \sum_{d | n} f(d) = \sum_{d \le x} f(d) \left\lfloor \frac{x}{d} \right\rfloor = \sum_{q \le x} F\left(\frac{x}{q}\right)$$

- We also obtain certain weighted sums of μ and Λ:

$$\sum_{n \le x} \mu(n) \left\lfloor \frac{x}{n} \right\rfloor = 1;$$

$$\sum_{n \le x} \Lambda(n) \left\lfloor \frac{x}{n} \right\rfloor = \log \lfloor x \rfloor ! = \sum_{n \le x} \log n = x \log x - x + O(\log x)$$

Example 1.5.2 (Exercise 3.6; [Apo98])

For $x \geq 2$, prove that

$$\sum_{n\leq x} \frac{\varphi(n)}{n^2} = \frac{1}{\zeta(2)} \log x + \frac{C}{\zeta(2)} - A + O\left(\frac{\log x}{x}\right)$$

where $A = \sum_{n=1}^{\infty} \frac{\mu(n)\log n}{n^2}$.

Proof.

$$\begin{aligned}
\sum_{n\leq x} \frac{\varphi(n)}{n^2} &= \sum_{n\leq x} \frac{1}{n} \sum_{d|n} \frac{\mu(d)}{d} \\
&= \sum_{d\leq x} \sum_{q\leq x/d} \frac{\mu(d)}{d^2 q} \\
&= \sum_{d\leq x} \frac{\mu(d)}{d^2} \sum_{q\leq x/d} \frac{1}{q} \\
&= \sum_{d\leq x} \frac{\mu(d)}{d^2} \left\{\log\left(\frac{x}{d}\right) + C + O\left(\frac{1}{x}\right)\right\} \\
&= \frac{1}{\zeta(2)} \log x + \frac{C}{\zeta(2)} - \sum_{d\leq x} \frac{\mu(d)\log d}{d^2} + O\left(\frac{1}{x}\right) \\
&= \frac{1}{\zeta(2)} \log x + \frac{C}{\zeta(2)} - \sum_{d=2}^{\infty} \frac{\mu(d)\log d}{d^2} + \sum_{d>x} \frac{\mu(d)\log d}{d^2} + O\left(\frac{1}{x}\right) \\
&= \frac{1}{\zeta(2)} \log x + \frac{C}{\zeta(2)} - A + O\left(\frac{\log x}{x}\right).
\end{aligned}$$

□

Theorem 1.5.3

For $x \geq 1$, we have

$$\left|\sum_{n\leq x} \frac{\mu(n)}{n}\right| \leq 1.$$

Equality holds only for $x < 2$.

Proof. For $x < 2$, clearly the equality holds, since $\mu(1) = 1$. For $x \geq 2$,

$$1 = \sum_{n\leq x} \mu(n) \left\lfloor \frac{x}{n} \right\rfloor = x \sum_{n\leq x} \frac{\mu(n)}{n} - \sum_{n\leq x} \mu(n) \left\{\frac{x}{n}\right\}$$

which implies,

$$\begin{aligned}
\left| x \sum_{n\leq x} \frac{\mu(n)}{n} \right| &= \left| 1 + \sum_{n\leq x} \mu(n) \left\{\frac{x}{n}\right\} \right| \\
&\leq 1 + \{x\} + \left| \sum_{2\leq n\leq x} \mu(n) \left\{\frac{x}{n}\right\} \right| \\
&< 1 + \{x\} + \lfloor x \rfloor - 1 = x
\end{aligned}$$

from which the result follows as desired. □

Remark 1.5.4. As a consequence, the partial sum involved in the above expression converges. In fact, it converges to 0, which is enough to prove Prime Number theorem (we will see this equivalence later!).

Let's look at a small problem:

Problem 1.5.5. For $x \geq 2$, prove that

$$\sum_{n \leq x} \frac{\mu(n)}{n} \left\lfloor \frac{x}{n} \right\rfloor = \frac{x}{\zeta(2)} + O(\log x)$$

Solution. For $x \geq 2$, we have

$$\begin{aligned} \sum_{n \leq x} \frac{\mu(n)}{n} \left\lfloor \frac{x}{n} \right\rfloor &= \sum_{n \leq x} \frac{\mu(n)}{n} \left(\frac{x}{n} - \left\{ \frac{x}{n} \right\} \right) \\ &= \frac{x}{\zeta(2)} - x \sum_{n > x} \frac{\mu(n)}{n^2} + O\left(\sum_{n \leq x} \frac{1}{n} \right) \\ &= \frac{x}{\zeta(2)} + x \cdot O\left(\frac{1}{x} \right) + O(\log x) \\ &= \frac{x}{\zeta(2)} + O(\log x). \end{aligned}$$

□

Theorem 1.5.6 (Dirichlet's Hyperbola Method)

Let f, g be Arithmetic Functions. Suppose $h = f * g$ and

$$F(x) = \sum_{n \leq x} f(n); \quad G(x) = \sum_{n \leq x} g(n); \quad H(x) = \sum_{n \leq x} h(n).$$

Choose any two reals a, b such that $ab = x$. Then

$$H(x) = \sum_{n \leq a} f(n) G\left(\frac{x}{n} \right) + \sum_{n \leq b} g(n) F\left(\frac{x}{n} \right) - F(a)G(b).$$

Proof. Observe that,

$$H(x) = \sum_{\substack{d,q \\ dq \leq x}} f(d)g(q).$$

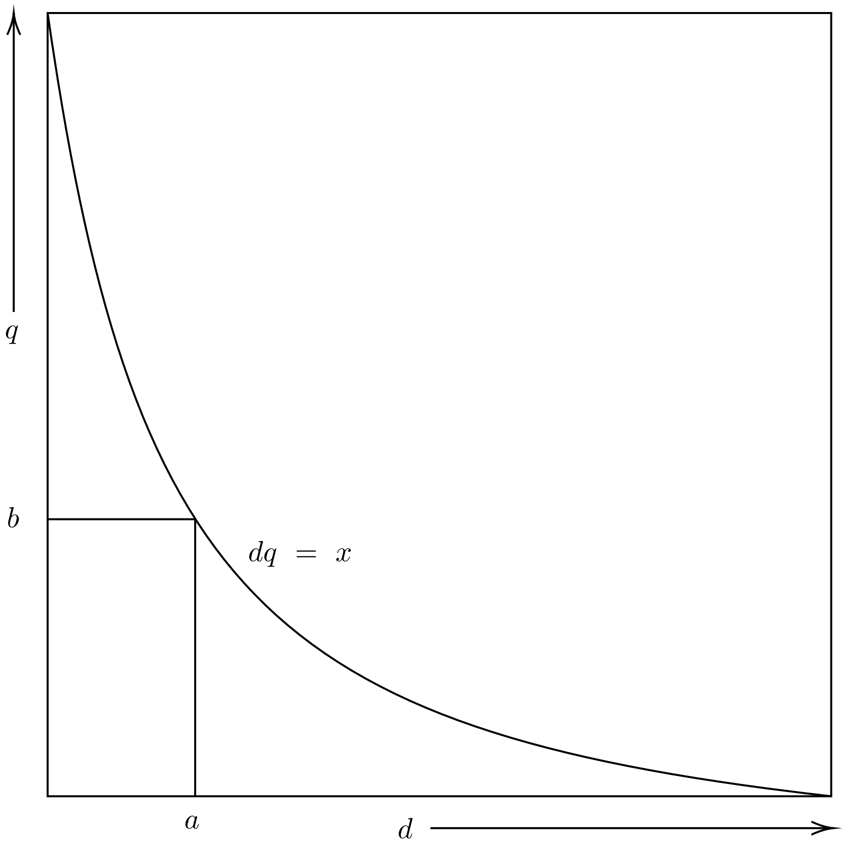
We can re-write the index set as union of two sets A, B, where

$$A = \{(d,q) : d \leq a, q \leq x/d\}; \quad B = \{(d,q) : q \leq b, q \leq x/q\}.$$

But there's an overlap between A and B:

$$A \cap B = \{(d,q) : d \leq a \text{ and } q \leq b\}.$$

A supportive figure is shown:



Our index set is the lattice points which fall exactly under the hyperbola $dq = x$. Therefore, breaking the sum in the expression of $H(x)$ into sums over two sets A and B followed by a subtraction of common elements, gives the desired result. □

2 Some Equivalent Forms of PNT

A significant milestone of number theory is the Prime Number Theorem (PNT). Let's assert it.

Theorem 2.0.1 (Prime Number Theorem)
If $\pi(x)$ denotes the number of primes $\leq x$, then

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} = 1$$

i.e.,

$$\pi(x) \sim \frac{x}{\log x}.$$

Gauss (1792) and Legendre (1798), independently conjectured it by careful inspection of tables of values. The conjecture was first proved, hundred years later, in 1896, independently by Jacques Hadamard and de la Vallee Poussin. They developed the proof via techniques of complex analysis, boiling the problem down to analyzing zeroes of Riemann's zeta function. However, the curiosity of proving the theorem in "elementary" means (i.e., bypassing complex analysis), was still on, although many great mathematicians started to believe such elementary proof can't possibly exist. Almost half a century later, again almost independently, Atle Selberg and Paul Erdos managed to prove the prime number theorem in elementary manner. Both of these proofs begin with considering a *fundamental* asymptotic formula, which was proved by Selberg in 1948. We will prove that formula in this chapter.

Another important perspective is that, there are couple of problems which are equivalent to PNT, and proving them is equally hard. We will discuss about these problems.

§2.1 Useful Definitions

Definition 2.1.1
For $x > 0$, the prime counting function is defined as:

$$\pi(x) = \sum_{n\leq x} a(n) = \sum_{p\leq x} 1$$

where, $a(n)$ is prime-characteristic function (also denoted as $1_{\mathcal{P}}$), defined as:

$$a(n) = \begin{cases} 1 & \text{if } n \text{ is prime,} \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2.1.2 (von Mangoldt's Function)
Define an arithmetic function Λ, such that

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^m,\ p \text{ is prime} \\ 0 & \text{otherwise.} \end{cases}$$

We illustrate some properties of it:

- Λ is not multiplicative.

- $\Lambda * 1 = \log$, and thus $\Lambda = \mu * \log$, via Theorem 1.1.5.

Definition 2.1.3

For $x > 0$, define Chebyshev's θ-function:

$$\theta(x) = \sum_{n \le x} \Lambda_1(n) = \sum_{n \le x} a(n) \log n = \sum_{p \le x} \log p$$

where, $\Lambda_1(n)$ is *tweaked* Mangoldt's function, defined as:

$$\Lambda_1(n) = \begin{cases} \log n & \text{if } n \text{ is prime} \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2.1.4

For $x > 0$, define Chebyshev's ψ-function:

$$\psi(x) = \sum_{n \le x} \Lambda(n) = \sum_{p \le x} \sum_{\substack{m=1 \\ p^m \le x}}^{\infty} \log p = \sum_{m \le \log_2 x} \sum_{p \le x^{1/m}} \log p = \sum_{m \le \log_2 x} \theta(x^{1/m})$$

The second last equality follows from an interchange of sums followed by the argument: for $x^{1/m} < 2$, no prime $p \le x^{1/m}$ will appear in the index of the sum.

§2.2 PNT and Chebyshev's Functions

We have already seen in the last chapter estimating partial sums of various arithmetical functions. It turns out that, an asymptotic estimate of partial sum of the von Mangoldt's function is equivalent to PNT.

Theorem 2.2.1

The following statements are equivalent:

1. $\pi(x) \sim \frac{x}{\log x}$
2. $\psi(x) \sim x$
3. $\theta(x) \sim x$.

Let's explore a few results and prove the above equivalences along the way.

Lemma 2.2.2

For $x > 0$, we have

$$0 \le \frac{\psi(x)}{x} - \frac{\theta(x)}{x} \le \frac{(\log x)^2}{2\sqrt{2}\log 2}$$

Remark 2.2.3. Immediately, we obtain (2) $\iff$ (3) of Theorem 2.2.1.

Proof. Left inequality is clear, by definition. For the right one, we plug the trivial bound $\theta(x) \le \sum_{p \le x} \log x \le x \log x$, into the required difference. □

Lemma 2.2.4 (Abel's Lemma)
For any arithmetic function $a(n)$, suppose

$$A(x) = \sum_{n \leq x} a(n).$$

Let f be a real-valued function with continuous derivative on $[y, x]$. Then we have

$$\sum_{y<n\leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)dt.$$

Proof. In terms of Riemann-Stieltjes integration,

$$\sum_{y<n\leq x} a(n)f(n) = \sum_y^x f(t)dA(t).$$

Integration by parts yields the desired result. □

Using Abel's lemma, we can find out switch back and forth between $\theta(x)$ and $\pi(x)$.

Theorem 2.2.5
For $x \geq 2$, we have

$$\theta(x) = \pi(x)\log x - \int_2^x \frac{\pi(t)}{t} dt$$

and

$$\pi(x) = \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t\log^2 t} dt.$$

Proof. Use Abel's lemma (specifically on $[3/2, x]$, while writing $\pi(x)$ in terms of $\theta(x)$). □

Theorem 2.2.6
The following statements are equivalent:

1. $$\lim_{x\to\infty} \frac{\pi(x)\log x}{x} = 1$$

2. $$\lim_{x\to\infty} \frac{\pi(x)\log \pi(x)}{x} = 1$$

3. $$\lim_{n\to\infty} \frac{p_n}{n\log n} = 1$$

where, p_n is n-th prime.

Proof. For (1) to (2), we take logarithms on (1) and deduce

$$\lim_{x\to\infty} \frac{\log \pi(x)}{\log x} = 1.$$

Taking product with (1), yields (2). Next, we assume (2) holds. If $x = p_n$, then $\pi(x) = n$ and

$$\pi(x)\log \pi(x) = n\log n.$$

Therefore, (3) follows. Next, assume (3) holds. We shall show (2). Given x, define n such that $p_n \leq x < p_{n+1}$, so that $n = \pi(x)$. It follows that

$$\frac{p_n}{n \log n} \leq \frac{x}{n \log n} < \frac{p_{n+1}}{n \log n} = \frac{p_{n+1}}{(n+1)\log(n+1)} \frac{(n+1)\log(n+1)}{n \log n}$$

Taking limits, we obtain

$$\lim_{n\to\infty} \frac{x}{n \log n} = 1$$

which implies (2). For (2) to (1), we take logarithms on (2), we deduce

$$\lim_{x\to\infty} \frac{\log x}{\log \pi(x)} = 1.$$

Together with (2), we obtain (1). □

§2.3 Shapiro's Tauberian Theorem

Tauberian theorems in general relate different weighted sums of arithmetic functions (sequences in terms of analysis). Shapiro proposed this theorem in 1950. It can be used to estimate different weighted sums if a specific one is known.

Theorem 2.3.1

Let $\{a(n)\}_{n=1}^{\infty}$ be a sequence of non-negative reals, such that $\forall x \geq 1$,

$$\sum_{n\leq x} a(n) \left[\frac{x}{n}\right] = x \log x + O(x).$$

Then the following statements hold:

1. For $x \geq 1$, we have
$$\sum_{n\leq x} \frac{a(n)}{n} = \log x + O(1).$$
(as if the floor in the hypothesis was cleaned!)
2. $\exists c_1 > 0$, such that $\forall x \geq 1$, we have
$$\sum_{n\leq x} a(n) \leq c_1 x.$$
3. $\exists c_2 > 0$ and $x_0 > 0$, such that $\forall x \geq x_0$, we have
$$\sum_{n\leq x} a(n) \geq c_2 x$$

For the proof, we refer to [Apo98]. We will look at some applications rather. This will make a clear picture of how powerful the theorem is.

Beautiful Applications

Previously we had stated that,

$$\sum_{n\leq x} \Lambda(n) \left[\frac{x}{n}\right] = x \log x + O(x).$$

Therefore, applying the Tauberian theorem, we get

$$\sum_{n\leq x} \frac{\Lambda(n)}{n} = \log x + O(1).$$

along with the following:

- $\exists c_1 > 0$, such that $\forall x \geq 1$, we have
$$\psi(x) \leq c_1 x.$$
In other words, this says, $\psi(x) = O(x)$.

- $\exists c_2 > 0$ and $x_0 > 0$, such that $\forall x \geq x_0$, we have
$$\psi(x) \geq c_2 x$$

We can also estimate,

$$\sum_{p \leq x} \log p \left[\frac{x}{p}\right] = \sum_{n \leq x} \Lambda_1(n) \left[\frac{x}{n}\right] = \sum_{n \leq x} \Lambda(n) \left[\frac{x}{n}\right] + O(x) = x \log x + O(x).$$

Therefore, applying the Tauberian theorem, we get

$$\sum_{n \leq x} \frac{\Lambda_1(n)}{n} = \sum_{p \leq x} \frac{\log p}{p} = \log x + O(1).$$

along with the following:

- $\exists c_1 > 0$, such that $\forall x \geq 1$, we have
$$\theta(x) \leq c_1 x.$$
In other words, we have $\theta(x) = O(x)$.

- $\exists c_2 > 0$ and $x_0 > 0$, such that $\forall x \geq x_0$, we have
$$\theta(x) \geq c_2 x$$

Remark 2.3.2. The asymptotic formula 2.3 for partial sum of $\log p/p$, is in fact, the starting point of Shapiro's proof to Dirichlet's theorem. Another application is to obtain an asymptotic formula for partial sum of $1/p$.

We have seen how Chebyshev's functions are naturally connected to PNT. In the upcoming sections, we will investigate whether different weighted sums of Mobius function is connected to PNT.

§2.4 PNT and Sums of Mobius function

Consider the following partial sum of Mobius function:

$$M(x) = \sum_{n \leq x} \mu(n).$$

In the previous chapter, we had seen that $|M(x)| \leq 1$, for all $x \geq 1$. It is possible that there are more or less equal numbers of square-free numbers which are built up with odd and even number of prime factors. In other words, we may expect almost equal occurrences of $\mu(n) = 1$ and $\mu(n) = -1$, when $n \leq x$ and x being large. In that case, $M(x) \to 0$ as $x \to \infty$. Won't it be astonishing if we manage to show that such behavior of $M(x)$ is equivalent to prime number theorem?

We will of course prove that, but in a couple of steps. First we show as $x \to \infty$, $\psi(x) \sim x$ implies $M(x) = o(x)$, via another function which is closely related to both μ and ψ.

Lemma 2.4.1

Define $H(x) = \sum_{n \leq x} \mu(n) \log n$. Then

$$H(x) = o(x \log x) \iff M(x) = o(x).$$

Proof. By Lemma 2.2.4, for $x \geq 1$:

$$H(x) = M(x) \log x - \int_1^x \frac{M(t)}{t}\, dt,$$

which implies,

$$\left| \frac{H(x)}{x \log x} - \frac{M(x)}{x} \right| \leq \frac{1}{x \log x} \left| \int_1^x \frac{M(t)}{t}\, dt \right| \leq \frac{1}{x \log x} \int_1^x \frac{|M(t)|}{t}\, dt \leq \frac{1}{\log x}.$$

This completes the proof. □

Theorem 2.4.2

We have the following implication:

$$\psi(x) = x + o(x) \implies M(x) = o(x), \quad \text{as } x \to \infty.$$

Proof. Assume that,

$$\psi(x) = x + o(x) \quad \text{as } x \to \infty.$$

That is, for a given $\varepsilon > 0$, there exists $x_0 > 0$ such that for all $x \geq x_0$, we have

$$|\psi(x) - x| < \varepsilon x.$$

It suffices to show, $\lim_{x\to\infty} \frac{H(x)}{x \log x} = 0$. Notice that,

$$\Lambda = \mu * \log = -\mu \log * 1,$$

which implies,

$$\mu \log = -\mu * \Lambda.$$

Then for $x \geq 1$,

$$-H(x) = -\sum_{n \leq x} \mu(n) \log n = \sum_{d \leq x} \mu(d) \psi\left(\frac{x}{d}\right).$$

To estimate right hand side, we need to break the sum over two intervals. The first, where d is such that $\psi\left(\frac{x}{d}\right) - \frac{x}{d}$ is small. The other, where applying the bound $\psi(x) = O(x)$, won't cause any harm. If $y = \frac{x}{x_0}$, we have

$$-H(x) = \sum_{d \leq y} \mu(d) \left\{ \psi\left(\frac{x}{d}\right) - \frac{x}{d} \right\} + x \sum_{d \leq y} \frac{\mu(d)}{d} + \sum_{y \leq d \leq x} \mu(d) \psi\left(\frac{x}{d}\right).$$

Therefore, for some constants $A, B > 0$ (independent of ε),

$$|H(x)| \leq \varepsilon \sum_{d \leq y} \frac{1}{d} + x + x \sum_{y \leq d \leq x} \frac{1}{d} < \varepsilon(1 + \log x) + x + x(\log x - \log \frac{x}{x_0} + B) = \varepsilon(1 + \log x) + x + Ax.$$

thus,

$$\left| \frac{H(x)}{x \log x} \right| < \frac{\varepsilon(1 + \log x) + x + Ax}{x \log x},$$

which can be made smaller than 4ε, for sufficiently large x. This completely the proof. □

Next, we are going to prove the converse.

Theorem 2.4.3

We have

$$M(x) = o(x) \implies \psi(x) = x + o(x), \quad \text{as } x \to \infty.$$

§2.5 Selberg's Asymptotic Formula

II

Dirichlet Characters

3 Primes in Arithmetic Progressions: An Introduction

One of the central question in number theory is about the distribution of prime numbers in an arithmetic progression.

§3.1 A Few Results and Problems

Theorem 3.1.1 (cf. Sierpinski)
If there is $h, k > 0$ and $n > 1$ such that

$$h, h+k, \ldots, h+(n-1)k$$

are all odd primes, then the common difference k must be divisible by each of the primes $< n$.

Proof. Let's notice the following:

- Observe that, $h \geq n$. Otherwise, if $h < n$, then $h + hk = h(k+1)$ would not be a prime, unless $h = 1$, which is again not a prime.

- Consider a prime $p < n \leq h$. Since, each of the numbers $h, h+k, \ldots, h+(p-1)k$ are primes, p cannot divide any of the number.

- Then, for those p numbers when divided by p, at least two of the remainders obtained must be same. This is because only $p-1$ remainders are available for p numbers; Pigeon-Hole principle.

- If those two numbers are $h+ik$ and $h+jk$ with $i > j$, then $p|(i-j)k$. But, $0 < i-j \leq p-1$ forces p to divide k, as required.

□

Remark 3.1.2. If $h > n$, then the same proof works to show that k must be divisible by all the primes $\leq n$. Note the inclusion of the possibility of $n|k$ if n is prime.

Corollary 3.1.3
If there is an increasing arithmetic progression of length $n > 2$ consisting only of odd primes, then the common difference must be divisible by P_n, the product of all the primes $< n$. Thus, it is at least P_n.

Remark 3.1.4. Embedding the previous remark into the corollary, if there is an AP with the first term $h > n$, the length, then the common difference must be divisible by P'_n, the product of all the primes $\leq n$. Thus, it is at least P'_n. It matters only when n is prime.

Example 3.1.5

Given an arithmetic progression of integers

$$h, h+k, \ldots, h+nk, \ldots$$

where $0 < k < 2000$. If $h + nk$ is prime for $n = t, t+1, \ldots, t+r$, then show that $r \leq 9$. In other words, at most 10 consecutive terms of this progression can be primes.

Solution. Observe that,

- If at least 12 consecutive terms of some progression with common difference k are primes, then all the primes < 12, namely $2, 3, 5, 7, 11$ must be dividing k. So, $k \geq 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 = 2310 > 2000$. Therefore, at most 11 consecutive primes can be present with the given $k < 2000$.

- Note that, if $t > 0$, then the first term $h + tk > h \geq 11$. By the remark above, again $k \geq 2310$.

- We are left with $t = 0$ case, i.e., to check whether $11, 11+k, \cdots, 11+10k$ is prime, when $210 \leq k < 2000$. This can be done by a SageMath program.

```
for k in range(210, 2000):
    A = [11 + q*k for q in range(11) if (11+q*k).is_prime()]
    B.append(len(A)==11)
```

  It turns out that the list B just contains “False" elements. Therefore, there are no such progression having 11 primes in it.

□

Let's note a few facts on primes in arithmetical progressions.

1. There exists precisely one arithmetical progression consisting of prime numbers whose common difference is 2, namely 3, 5, 7. If there is other 3-length AP, starting with > 3, then 3 must divide the common difference 2, which is impossible (follows from the remark).

2. There are infinitely many 3-length APs (with different first term or common difference) consisting only of 3 primes. Examples are: $\{3, 7, 11\}$, $\{3, 11, 19\}$, $\{3, 43, 83\}$.

3. The common difference of an arithmetical progression consisting of a hundred prime numbers would have to be divisible by the product of all prime numbers less than 100, and thus it would have more than thirty digits!

Conjecture 3.1.6 (Erdos, 1976)**.** If $\{\alpha_n\}$ is a sequence of positive integers, such that the series

$$\sum_{n=1}^{\infty} \frac{1}{\alpha_n}$$

diverges, then for every $m \in \mathbb{N}$, the sequence $\{\alpha_n\}$ contains an AP of length m.

However, some special cases have been proved very recently:

Theorem 3.1.7 (Green-Tao, 2004)

The prime numbers contain arbitrarily long arithmetic progressions.

Theorem 3.1.8 (Bloom-Sisak, 2020)

If $\{\alpha_n\}$ is a sequence of positive integers, such that the series

$$\sum_{n=1}^{\infty} \frac{1}{\alpha_n}$$

diverges, then the sequence $\{\alpha_n\}$ contains an AP of length 3.

Next we move to a question of different taste: Given a common difference $k > 0$ and first term $h > 0$, does there exist infinitely many primes in the arithmetic progression $\{h + kn\}_{n=1}^{\infty}$? If $(h, k) = d > 1$, then clearly, answer is no. Therefore, $(h, k) = 1$ is necessary in order to have existence of infinitely many primes. Dirichlet in 1837, showed that, it is sufficient as well.

Theorem 3.1.9 (Dirichlet, 1837)

If $(h, k) = 1$, then there exists infinitely many primes of the form $h + kn$.

We will discuss about a proof of the theorem, given by Shapiro (1950), in the next chapter. It involves a technique from character theory on finite abelian groups.

Theorem 3.1.10

The following statements are equivalent:

1. If $(h, k) = 1$, then there exists infinitely many primes of the form $h + kn$.
2. If $(h, k) = 1$, then there exists at least one prime of the form $h + kn$.

Proof. (1) to (2) is trivial. For the converse, if $k = 1$, the theorem is already true, since there are infinitely many primes. So we safely assume, $k > 1$. Let $m \in \mathbb{N}$. Since $(h, k) = 1$ we have $(h, k^m) = 1$. By (2), there is a prime p such that $p = h + k^m n$ for some $n \in \mathbb{N}$. But $p = h + k^m n > k^m n \geq 2^m > m$. We have thus shown that, for any $m \in \mathbb{N}$, there is a prime p of the form $h + kn$, which is bigger than m. This completes the proof. □

§3.2 Hunting Co-primes in Arithmetic Progressions

In this section, we describe a few results which will be used later. The proofs generally require similar constructions.

Theorem 3.2.1

Let a, d be positive integers, such that $(a, d) = 1$. For every integer $k \geq 1$, there is an integer m such that $(m, k) = 1$ and $m = a + qd$ for some integer q. In other words, for each postive integer k we can always find an integer relatively prime to k lying in an arithmetic progression formed by first term a and common difference d.

Proof. Let's go through the key idea:

- *Key Idea:* Observe that, a and d are already coprime. If we choose a prime p such that p does not divide a, then $(a, pd) = 1$. Note that p cannot divide m. Moreover, if this prime p divides k, then we can expect m and k to be co-prime.

Suppose, $p_1, \ldots, p_n$ be distinct primes such that $p_i | k$ and $p_i \not| a$. Set $q = p_1 \ldots p_n$. We claim that, if $m = a + qd$, then $(m, k) = 1$. If not, then there is a prime $p | m$ and $p | k$.

- If $p|a$, then $p|qd$. Now, $p \not| q$, because if $p|q$, then together with $p|k$ and $p|a$, it would contradict the definition of q. This means, $p|d$, which again contradicts the fact $(a,d)=1$.
- Else, $p \not| a$. Then $p|q$, by definition. Since, $p|m$, we have $p|a$, which is a contradiction.

□

Remark 3.2.2. The proof provides an explicit construction of an integer in the above AP, co-prime to a given integer.

Proof. (Alternate) Assuming Dirichlet's theorem (which we will prove later), there are infinitely many primes in the arithmetic progression $\{a+qd\}_{q=0}^{\infty}$, when $(a,d)=1$. For any positive integer $k \geq 1$, there are only finitely many primes dividing k. Choose a prime from the AP, which is not a divisor of k. Call that m, so that $(m,k)=1$. □

§3.3 Weaker Dirichlet's Theorem via Cyclotomic Polynomials

4 Fourier Analysis on Finite Abelian Groups

§4.1 Finite Fourier Series

Consider the set of functions from a finite abelian group G to $\mathbb{C}$ over the field $\mathbb{C}$. Endowed with pointwise addition and scalar multiplication it forms a vector space. Let a_is (for $i \in \{1, \dots, |G|\}$) denote the members of G. Then, by the following isomorphism:

$$f \mapsto (f(a_1), \dots, f(a_{|G|}))$$

we note that the space is $|G|$-dimensional. In the next step, we equip it with a Hermitian inner product:

$$(f, g) = \frac{1}{|G|} \sum_{a \in G} f(a)\overline{g(a)}$$

A typical heuristic in vector space theory is to construct a suitable basis, which would have *nice* properties. According to our context, these properties are listed below:

- The basis should orthogonal (can be normalized later, if required), with respect to the equipped inner product.
- The members of the basis should homomorphisms from G to $\mathbb{C}^\times$ (exponentials!).

Of course, the above list is motivated from the Fourier theory on complex-valued functions defined on circle (i.e., periodic functions defined on $\mathbb{R}$).

§4.1.1 Special Case: G = Z(k)

We set $G = \mathbb{Z}(k)$, i.e., the group of residues modulo k and try to find a basis of the vector space mentioned above. First for each $i \in \{0, \dots, k-1\}$, we define $e_i : \mathbb{Z}(k) \to \mathbb{C}$ such that

$$e_i(j) = \zeta^{ij}$$

where $\zeta = e^{2\pi i/k}$. Then let us consider the following set of k functions:

$$\widehat{\mathbb{Z}(k)} = \{e_0, e_1, \dots, e_{k-1}\}.$$

and observe two facts: They are homomorphisms from $\mathbb{Z}(k)$ to $\mathbb{C}^\times$ and that

$$(e_i, e_j) = \frac{1}{k} \sum_{l \in \mathbb{Z}(k)} \zeta^{(i-j)l} = \frac{1}{k} \sum_{l=0}^{k-1} \zeta^{(i-j)l} = \begin{cases} 1 & \text{if } i \equiv j \bmod k \\ 0 & \text{otherwise.} \end{cases}$$

Since $0 \le i, j \le k-1$, we must $i \equiv j \bmod k$ if and only if $i = j$. Therefore, the set is orthogonal. In fact, they are orthonormal (which we desired!) and linearly independent. Since, the space is k-dimensional $\widehat{\mathbb{Z}(k)}$ is indeed a orthonormal basis. Therefore, any complex valued function $f : \mathbb{Z}(k) \to \mathbb{C}$ can be written in terms of the above basis:

Theorem 4.1.1 (Finite Fourier Series)

Let $f : \mathbb{Z}(k) \to \mathbb{C}$. Then

$$f(m) = \sum_{e \in \widehat{\mathbb{Z}(k)}} (f, e)e(m)$$

where

$$(f, e) = \frac{1}{k} \sum_{a \in \mathbb{Z}(k)} f(a)\overline{e(a)}$$

Proof. If $f(m) = \sum_{e \in \widehat{\mathbb{Z}(k)}} a(e)e(m)$, then the coefficient $a(e)$ can be determined by taking inner product of both sides of the equation with e. Linearity of inner product and orthogonality of basis vectors lead to the required result. □

§4.1.2 General Case: Theory of Characters

Let G be a finite abelian group. Let us define the homomorphisms (which we want to be our basis).

Definition 4.1.2

A map $f : G \to \mathbb{C}^\times$ is said to be a *character* on G, if it is a group homomorphism from G to the multiplicative group of non-zero complex numbers.

For instance, the basis vectors we had when $G = \mathbb{Z}(k)$, were indeed characters on the additive group $\mathbb{Z}(k)$. In general setting, we will not be so fortunate to get explicit forms of the characters. But we can manage to show that they form a basis of the set of complex valued functions on G, outlined below:

1. There are $|G|$ distinct characters that can be defined on finite abelian group G.
2. These characters form another finite abelian group $\widehat{G}$, namely *dual* of G.
3. Orthogonality relations with respect to the defined inner product are deduced from (2) and properties of characters (in fact, characters are orthonormal).
4. Orthogonality will force linear independence. Therefore, characters form basis.

§4.2 Ramanujan Sums

Definition 4.2.1 (Ramanujan Sum)

Let $k, n > 0$. Then the Ramanujan sum $c_k(n)$ is defined as:

$$c_k(n) = \sum_{\substack{m \bmod k \\ (m,k)=1}} e^{\frac{2\pi i}{k}nm}$$

Srinivasa Ramanujan introduced these sums in a 1918 paper. They have some interesting properties. For instace, we have

- A special case: $c_k(1)$, the sum over all the primitive k-th roots of unity, is $\mu(k)$.
- Another one: $c_k(k)$ counts the number of primitive k-th roots of unity, which is $\varphi(k)$.
- The Ramanujan sum $c_k(n)$ can be expressed as "almost a convolution":

$$c_k(n) = \sum_{d|(n,k)} d\mu\left(\frac{k}{d}\right)$$

In general, let's consider two arithmetical functions f, g. Let $s_k(n)$ denote the following new type of convolution of f and g:

$$s_k(n) = \sum_{d|(n,k)} f(d)g\left(\frac{k}{d}\right)$$

Observe that, s_k is periodic modulo k. This follows, from the property of gcd. Thus, we expect a finite fourier expansion of s_k.

§4.3 Gauss Sums

5 Special Topic: Polya-Vinogradov Inequality

We have encountered Dirichlet characters. They are arithmetic functions and in fact, they are periodic. Thus, a natural question arises about investigating bounds of their partial sums ("character sum"). Let χ be a Dirichlet character modulo q.

§5.1 Principal Characters

Theorem 5.1.1

If $\chi = \chi_0$, the principal character modulo q, then for $x \geq 1$ we have

$$\sum_{m \leq x} \chi(m) = \frac{\varphi(q)}{q} x + O(1).$$

Proof. Let $x = qk + r$, with $0 \leq r < q$. Then

$$\sum_{m \leq x} \chi(m) = \sum_{s=1}^{k} \sum_{(s-1)q < m \leq sq} \chi(m) + \sum_{qk < m \leq qk+r} \chi(m) = \varphi(q)k + O(1) = \frac{\varphi(q)}{q} x + O(1).$$

□

The result is as expected because partial sum of the principal character $\leq x$ ultimately counts the number of positive integers $\leq x$. And in each of the $\approx \frac{x}{q}$ intervals, between two consecutive multiples of q, there are exactly $\varphi(q)$ such integers.

§5.2 Non-principal Characters

In this case, we can get a trivial bound:

Theorem 5.2.1

If χ is a non-principal character modulo q, then for $x \geq 1$ we have

$$\left| \sum_{m \leq x} \chi(m) \right| < \varphi(q).$$

Proof. We use the fact that, if χ is non-principal, then $\sum_{m \bmod q} \chi(m) = 0$. Let $x = qk + r$, with $0 \leq r < q$ as before. Notice that, we have

$$\sum_{m \leq x} \chi(m) = \sum_{s=1}^{k} \sum_{(s-1)q < m \leq sq} \chi(m) + \sum_{qk < m \leq qk+r} \chi(m) = \sum_{m \leq r} \chi(m).$$

Invoking trivial bound on the last expression, we obtain

$$\left| \sum_{m \leq x} \chi(m) \right| = \left| \sum_{m \leq r} \chi(m) \right| = \left| \sum_{\substack{m \leq r \\ (m,q)=1}} \chi(m) \right| < \varphi(q).$$

Reason of the strict inequality: if index set of the sum can contain at most $\varphi(q)$ elements. Furthermore, if it contains all the $\varphi(q)$ elements, then the sum would be 0, which is less than $\varphi(q)$. Otherwise, if index set contains at most $\varphi(q) - 1$ elements, then the sum would be $\leq \varphi(q) - 1$ which is again $\varphi(q)$. □

If we are more careful, we can obtain a better bound:

Theorem 5.2.2

If χ is a non-principal character modulo q, then for $x \geq 1$ we have

$$\left|\sum_{m\leq x}\chi(m)\right| \leq \frac{\varphi(q)}{2}.$$

Proof. Let $x = qk + r$, with $0 \leq r < q$ as before. Again we write:

$$\sum_{m\leq x}\chi(m) = \sum_{s=1}^{k}\sum_{(s-1)q<m\leq sq}\chi(m) + \sum_{qk<m\leq qk+r}\chi(m) = \sum_{m\leq r}\chi(m) = \sum_{\substack{m\leq r\\(m,q)=1}}\chi(m).$$

If between 1 and $r < q$, there are at most $\varphi(q)/2$ integers relatively prime to q, then we are done. Otherwise, we may assume there are more than $\varphi(q)/2$ integers relatively prime to q, between 1 and r. In that case, between $r+1$ and q, i.e., in the complement set, there must be at most $\varphi(q)/2$ such elements. Since χ is non-principal, we have

$$\left|\sum_{m\leq x}\chi(m)\right| = \left|\sum_{\substack{m\leq r\\(m,q)=1}}\chi(m)\right| = \left|\sum_{\substack{m\leq q\\(m,q)=1}}\chi(m) - \sum_{\substack{r<m\leq q\\(m,q)=1}}\chi(m)\right| = \left|\sum_{\substack{r<m\leq q\\(m,q)=1}}\chi(m)\right| \leq \frac{\varphi(q)}{2},$$

as required. □

§5.3 Finite Fourier Series

So far, we have not explicitly used periodicity of the Dirichlet characters modulo q, which allows one to decompose it in terms of characters on $\mathbb{Z}(q)$.

Lemma 5.3.1

Let χ be a Dirichlet character modulo q. Then we have the following finite Fourier series:

$$\chi(m) = \sum_{r \bmod q} a_q(r)e\left(\frac{rm}{q}\right)$$

where the Fourier coefficients are given by:

$$a_q(r) = \frac{1}{q}\sum_{s \bmod q}\chi(s)e\left(-\frac{rs}{q}\right) = \frac{1}{q}G(-r,\chi).$$

Proof. A direct application of Theorem 4.1.1. □

For primitive characters, Gauss sum $G(n,\chi)$ is separable for all $n \in \mathbb{N}$. Thus, the Fourier coefficients become:

$$a_q(r) = \frac{1}{q}G(-r,\chi) = \frac{1}{q}\overline{\chi(-r)}G(1,\chi)$$

Thus, we can estimate the Fourier coefficients,

$$|a_q(r)| = \begin{cases}\frac{1}{\sqrt{q}} & \text{if } (r,q)=1\\ 0 & \text{otherwise.}\end{cases}$$

§5.4 The Polya-Vinogradov Inequality

§5.4.1 Primitive Case

The foregoing discussion provides a way to investigate a much better bound for partial sum of primitive characters.

Theorem 5.4.1 (Polya-Vinogradov Inequality; Primitive Characters)
Let χ be a primitive Dirichlet character modulo $q > 1$. Then for $x \geq 1$, we have

$$\left| \sum_{m \leq x} \chi(m) \right| < \sqrt{q} \log q.$$

Remark 5.4.2. The above inequality also holds for any non-principal character, upto some positive constant, as we will prove later.

Proof. We write the partial sum

$$\sum_{m \leq x} \chi(m) = \sum_{m \leq x} \sum_{r=1}^{q} a_q(r) e\left(\frac{rm}{q}\right) = \sum_{r=1}^{q} a_q(r) \sum_{m \leq x} e\left(\frac{rm}{q}\right)$$

via Lemma 5.3.1 and thus,

$$\begin{aligned} \left| \sum_{m \leq x} \chi(m) \right| &= \left| \sum_{r=1}^{q} a_q(r) \sum_{m \leq x} e\left(\frac{rm}{q}\right) \right| \\ &\leq \sum_{r=1}^{q} |a_q(r)| \left| \sum_{m \leq x} e\left(\frac{rm}{q}\right) \right| \\ &\leq \frac{1}{\sqrt{q}} \sum_{\substack{r \leq q \\ (r,q)=1}} \left| \sum_{m \leq x} e\left(\frac{rm}{q}\right) \right| = \frac{1}{\sqrt{q}} \sum_{\substack{r \leq q \\ (r,q)=1}} |E(x,r)| \end{aligned}$$

where $E(x,r) = \sum_{m \leq x} e\left(\frac{rm}{q}\right)$. Note that,

$$E(x, q-r) = \sum_{m \leq x} e\left(\frac{(q-r)m}{q}\right) = \sum_{m \leq x} e\left(-\frac{rm}{q}\right) = \overline{E(x,r)}$$

and $(r,q) = 1$ if and only if $(q-r,q) = 1$. As a consequence, we can reduce the index set.

$$\left| \sum_{m \leq x} \chi(m) \right| \leq \frac{1}{\sqrt{q}} \sum_{\substack{r \leq q \\ (r,q)=1}} |E(x,r)| \leq \frac{2}{\sqrt{q}} \sum_{\substack{r \leq q/2 \\ (r,q)=1}} |E(x,r)|.$$

Moreover, no characters modulo $q = 2m$ can be primitive, as m would be an induced modulus for it. Thus,

- Either $4|q$, in which case we already have $q/2, q > 1$. So, the sum can be taken upto $\frac{q-1}{2}$.
- Or, q is odd. In that case, again the sum can be taken upto $\frac{q-1}{2}$.

We now have to find an estimate for the last expression on right hand side. Let $k = \lfloor x \rfloor$. Notice that,

$$\begin{aligned} E(x,r) &= \sum_{m\le x} e\left(\frac{rm}{q}\right) \\ &= e(r/q)\cdot\frac{e(rk/q)-1}{e(r/q)-1} \\ &= e(r(k+1)/2q)\cdot\frac{e(rk/2q)-e(-rk/2q)}{e(r/2q)-e(-r/2q)} \\ &= e\left(\frac{r(k+1)}{2q}\right)\cdot\frac{\sin(\pi rk/q)}{\sin(\pi r/q)}. \end{aligned}$$

Then by trivial bounds we have,

$$|E(x,r)| \le \frac{1}{\sin(\pi r/q)}$$

Since, $r \le (q-1)/2$, safely we can conclude $0 < \frac{\pi r}{q} < \frac{\pi}{2}$, which implies

$$\frac{2}{\pi} = \frac{\sin(\pi/2)}{\pi/2} < \frac{\sin(\pi r/q)}{\pi r/q} < 1.$$

Therefore,

$$|E(x,r)| \le \frac{1}{\sin(\pi r/q)} < \frac{1}{2}\cdot\frac{q}{r}$$

and hence,

$$\left|\sum_{m\le x}\chi(m)\right| \le \frac{2}{\sqrt{q}}\sum_{\substack{r\le q/2\\(r,q)=1}}|E(x,r)| \le \frac{2}{\sqrt{q}}\sum_{\substack{r\le(q-1)/2\\(r,q)=1}}|E(x,r)| < \sqrt{q}\sum_{r\le(q-1)/2}\frac{1}{r}.$$

To bound the right side, we may use convexity of $f(r) = \frac{1}{r}$ for $r > 1$, i.e.,

$$\frac{1}{r} < \int_{r-\frac{1}{2}}^{r+\frac{1}{2}}\frac{1}{t}\,dt = \log\left(\frac{2r+1}{2r-1}\right)$$

thus,

$$\left|\sum_{m\le x}\chi(m)\right| < \sqrt{q}(\log\,[2\left(\frac{q-1}{2}\right)+1] - \log 1) = \sqrt{q}\log q,$$

as desired. □

§5.4.2 General Case

We can generalize the above statement for any non-principal Dirichlet character. For the proof to work, we need a theorem, which is crucial for its own sake.

Theorem 5.4.3

For every $\epsilon > 0$, we have

$$d(n) \ll n^{\epsilon}.$$

The theorem simplifies lots of problems in number theory. However, the standard proof requires the prime number theorem (PNT). Although, it can be proved without using PNT, as illustrated in Apostol, Exercise 13.13. Now we state the main result of this chapter.

Theorem 5.4.4 (Polya-Vinogradov Inequality, 1918)

Let χ be a non-principal Dirichlet character modulo q. Then for $x \geq 1$, we have

$$\sum_{m \leq x} \chi(m) \ll \sqrt{q} \log q.$$

Proof. Since χ mod q is non-primitive, let $d|q$, $d < q$ be the conductor of χ. Then there is a unique primitive character $\chi^\star$ mod d such that for every positive integer n,

$$\chi(n) = \chi^\star(n)\chi_0(n)$$

where χ_0 is the principal character modulo q. Then we write:

$$\sum_{m \leq x} \chi(m) = \sum_{\substack{m \leq x \\ (m,q)=1}} \chi^\star(m) = \sum_{m \leq x} \chi^\star(m) \sum_{\substack{r|q \\ r|m}} \mu(r) = \sum_{r|q} \mu(r)\chi^\star(r) \sum_{s \leq x/r} \chi^\star(s).$$

We can then apply Theorem 5.4.1 to obtain the following bound:

$$\left| \sum_{m \leq x} \chi(m) \right| < \left| \sum_{r|q} \mu(r)\chi^\star(r) \right| \sqrt{d} \log d.$$

Observe that, since $\chi^\star$ and μ are both multiplicative,

$$\sum_{r|q} \mu(r)\chi^\star(r) = \prod_{p|q} (1 - \chi^\star(p))$$

Notice that, if $p|q$ and $p|d$, then $(d,p) > 1$, implying $\chi^\star(p) = 0$. Thus,

$$\prod_{p|q} (1 - \chi^\star(p)) = \prod_{\substack{p|q \\ p\nmid d}} (1 - \chi^\star(p))$$

Hence,

$$\left| \sum_{r|q} \mu(r)\chi^\star(r) \right| = \left| \prod_{\substack{p|q \\ p\nmid d}} (1 - \chi^\star(p)) \right| \leq \prod_{p|(q/d)} |1 - \chi^\star(p)| \leq 2^{\nu(q/d)}.$$

where $\nu(n)$ is the number of distinct prime factors of n. Recall that, $2^{\nu(n)} \leq d(n)$, the number of divisors of n. Furthermore, $d(n) \ll \sqrt{n}$, in particular; via Theorem .Therefore,

$$\left| \sum_{m \leq x} \chi(m) \right| < 2^{\nu(q/d)} \sqrt{d} \log d \leq d(q/d)\sqrt{d} \log d \ll \sqrt{\frac{q}{d}} \sqrt{d} \log d < \sqrt{q} \log q,$$

as desired. □

Bibliography

[Apo98] Tom M Apostol. *Introduction to analytic number theory.* Springer Science & Business Media, 1998.