

Methods in Non-Classical Number Theory

Submitted by -

Pallav Goyal
3rd year Undergraduate Student
Department of Mathematics and Statistics
Indian Institute of Technology Kanpur

KVPY Registration no. - SA-11110157

Mentored by -

Dr. B. Sury
Professor
Stat-Math Unit
Indian Statistical Institute Bangalore

Contents

1	Algebraic Number Theory - Quadratic fields	4
1.1	Complex Lattices	5
1.2	Ideal Class Groups	9
1.3	Dirichlet Class Number Formula	13
2	Analytic Number Theory	17
2.1	Dirichlet's Theorem on the infinitude of primes in arithmetic progressions	17
2.1.1	Dirichlet Characters	17
2.1.2	L-series	18
2.1.3	The theorem	22
2.2	Brun's Theorem on the convergence of the sum of reciprocals of twin primes	24
2.2.1	Brun's Simple Pure Sieve	24
2.2.2	The theorem	26

Introduction

The theory of numbers has fascinated human beings since times immemorial. Starting with Euclid's proof for the infinitude of prime numbers, huge advancements have been made in Number Theory. With time however, the techniques in classical number theory, having been used to their full capacity, needed to be supplemented by newer tools from other fields of mathematics. This led to the birth of non-classical number theory with the likes of Gauss, Legendre, Dirichlet and Kummer as its pioneers. In this report, I would like to give a (not at all comprehensive!) glimpse of the strength these tools offer to us.

We'll start with algebraic aspects of number theory in Section 1. This subject itself arose from an attempt to solve Fermat's equation but has become central to many aspects of mathematics. To have a complete discussion of the subject of algebraic number fields, one would need to develop a sizeable amount of commutative algebra. On the other hand, the special case of quadratic fields already shares several features with the general case but can be studied with a more focussed, narrower background. In particular, the study of quadratic fields is possible using properties of lattices in \mathbb{C} . Our discussion will lead us without much ado to the proof of the so-called Dirichlet Class Number Formula, which is a classic example of blending ingredients from varied fields to get a sensational potpourri. In Section 2, we move on to analytic aspects of number theory and prove two extremely elegant and important results demonstrating the power of analytic tools. In particular, we prove Brun's theorem that the sum of the reciprocals of "twin" primes (primes p for which $p + 2$ is also prime) is finite. In general, both the algebraic and the analytic aspects need to be combined to get powerful results like Dirichlet's theorem on primes in arithmetic progressions.

Most of the pre-requisites have been discussed and can be read up from the references. It is assumed that the reader has a good command over rigorous proofs and arguments, elementary number theory, some basic analysis and a



familiarity with the Big-Oh notation.



1 Algebraic Number Theory - Quadratic fields

Many classical number-theoretic problems (for example, solutions of Diophantine equations) depend on the study of algebraic number fields and certain of their subrings. If the subring has nice algebraic properties like unique factorization into its “prime” elements analogous to what happens in integers, the Diophantine problem can be solved by going into that realm. For instance, one can show that the integral solutions of the Diophantine equation $x^2 + 2 = y^3$ are $(x, y) = (\pm 5, 3)$ using the fact that the subring

$$\mathbb{Z}[\sqrt{-2}] = \{a + b\sqrt{-2} : a, b \in \mathbb{Z}\}$$

of the quadratic field $\mathbb{Q}(\sqrt{-2})$ has unique factorization. In general, such rings in algebraic number fields (another name for finite extension fields of \mathbb{Q}) may not have unique factorization. Nevertheless, if one talks about ideals and prime ideals in such rings instead of individual elements and prime elements, Kummer and Dedekind showed that the unique factorization can be re-captured in this more general sense. Every non-zero, proper ideal turns out to have a unique factorization in terms of prime ideals! Thus, one can define a group associated to the number field by looking at classes of ideals (with respect to a certain equivalence relation) and their product rather than individual elements of the ring. The amazing fact is that this associated group - called the class group of the number field - turns out to be finite. In particular, there is a “finite obstruction” from unique factorization holding in these rings. Gauss had already developed some of these ideas in the case of subrings of quadratic fields (fields obtained by attaching the square-root of a - positive or negative - square-free integer to \mathbb{Q}) through their relation with binary quadratic forms. We study the algebraic number theory of imaginary quadratic fields. The study is facilitated by looking at complex lattices. Following a discussion of the finiteness of the “class number”, we’ll look at a proof of the celebrated Dirichlet Class Number Formula. The first few subsections will develop the crux of the theory that will be needed for the final proof.

1.1 Complex Lattices

Definition 1.1. A complex lattice Λ is a subset of \mathbb{C} such that:

$$\Lambda = \{m\alpha + n\beta : m, n \in \mathbb{Z}\}$$

where, α and β are two linearly independent complex numbers.

(α, β) here is called a basis of Λ and we say $\Lambda = \langle \alpha, \beta \rangle$. The basis is said to be normalised if $\text{Im}(\beta/\alpha) > 0$. It is easy to see that a complex lattice need not have a unique basis. In fact, we have the elementary result:

Lemma 1.2. Given a normalised basis $\{\alpha, \beta\}$ for a complex lattice Λ , and any $a, b, c, d \in \mathbb{Z}$ for which $ad - bc = \pm 1$, we have that $(a\alpha + b\beta, c\alpha + d\beta)$ is also a basis for Λ , and every basis is of this form.

Now, due to an abundance of bases, in order to identify a lattice using its bases, we define:

Definition 1.3. For a complex lattice Λ , we define its \mathcal{J} -set as:

$$\mathcal{J}(\Lambda) = \left\{ \frac{\beta}{\alpha} : (\alpha, \beta) \text{ is a normalised basis of } \Lambda \right\}$$

It is easy to verify that the \mathcal{J} -sets create a partition of the set of all complex numbers. Using this definition, we can define an equivalence relation (called homothety) between complex lattices: $\Lambda \sim \Lambda'$ iff $\mathcal{J}(\Lambda) = \mathcal{J}(\Lambda')$. Under this equivalence, it is seen that 2 complex lattices are homothetic if and only if one can be expressed as a constant complex number multiple of the other, i.e.,

$$\Lambda \sim \Lambda' \text{ if and only if } \exists \gamma \in \mathbb{C} \text{ such that } \Lambda = \gamma \cdot \Lambda'.$$

As the \mathcal{J} -sets partition the set of complex numbers, if it can be shown that 2 complex lattices have just one common basis, then they will be homothetic. In order to utilise this fact to classify homothetic lattices, we define:

$$\mathcal{F} = \left\{ z \in \mathbb{C} : \text{Im}(z) > 0, |z| \geq 1, -\frac{1}{2} < \text{Re}(z) \leq \frac{1}{2} \text{ and } \text{Re}(z) > 0 \text{ if } |z| = 1 \right\}$$

Now given $\mathcal{J}(\Lambda)$, it is a simple exercise to note that $\mathcal{F} \cap \mathcal{J}(\Lambda)$ consists

of exactly 1 element. We choose this element as a representative of the entire class of homothetic complex lattices with a given \mathcal{J} -set, and call it the j -invariant of the class of lattices.

Having classified all complex lattices upto homothety, we now look upon a very important operation we can perform on these lattices.

Definition 1.4. *We say that a complex lattice Λ has complex multiplication (or CM) by γ , if $\gamma.\Lambda$ is a sublattice of Λ*

In the above definition, we implicitly assume that γ is not an integer. Now, we state some results that can be derived using elementary algebra and so, their proofs have been omitted.

Theorem 1.5. *A complex lattice Λ can have CM by γ if and only if γ is of the form:*

$$\sqrt{-n}$$

if $n \equiv 1, 2 \pmod{4}$ or of the form

$$\frac{1 + \sqrt{-n}}{2}$$

otherwise, where n is a square free positive integer.

Theorem 1.6. *Let $n \equiv 1, 2 \pmod{4}$ be a square free positive integer. Then, a complex lattice having CM by $\sqrt{-n}$ must be homothetic to a complex lattice of the form:*

$$\left\langle 1, \frac{a + \sqrt{-n}}{b} \right\rangle$$

where:

- $a, b \in \mathbb{Z}$
- $0 < b < 2\sqrt{\frac{n}{3}}$
- $-b < 2a \leq b$
- $a^2 + n \geq b^2$ and $a \geq 0$ if $a^2 + n = b^2$
- $b|a^2 + n$

Theorem 1.7. *Let $n \equiv 3 \pmod{4}$ be a square free positive integer. Then, a complex lattice having CM by $\frac{1+\sqrt{-n}}{2}$ must be homothetic to a complex lattice of the form:*

$$\langle 1, \frac{a + \sqrt{-n}}{b} \rangle$$

where:

- $a, b \in \mathbb{Z}, a$ is odd, b is even
- $0 < b < 2\sqrt{\frac{n}{3}}$
- $-b < 2a \leq b$
- $a^2 + n \leq b^2$ and $a \geq 0$ if $a^2 + n = b^2$
- $2b \mid a^2 + n$

With the help of these theorems, for a given square free integer n , we define $\mathcal{Cl}(-n)$ to be the set of complex numbers satisfying the conditions given in Theorems 1.6 and 1.7. As this set is finite, we can define the class number $h(-n)$ to be the cardinality of this set.

Before ending with complex lattices, we prove a result (which shall be used later) approximating number of lattice points in a complex lattice for given norm range. We define:

$$C_t = \{z \in \mathbb{C} : |z| \leq t\}$$

and also, A to be the area of the parallelogram P with sides α and β where $\Lambda = \langle \alpha, \beta \rangle$.

Lemma 1.8. *Given a complex lattice Λ , there exists a constant C such that $\forall t$,*

$$|\#\Lambda \cap C_t - \frac{\pi t^2}{A}| \leq C.t$$

Proof. For each $\lambda \in \Lambda$, denote by P_λ the parallelogram with vertices: $\lambda, \lambda + \alpha, \lambda + \beta, \lambda + \alpha + \beta$. Next, we use the notation:

$$\begin{aligned}
n(t) &= \#(\Lambda \cap C_t) \\
n_1(t) &= \#\{\lambda : P_\lambda \subseteq C_t\} \\
n_2(t) &= \#\{\lambda : P_\lambda \text{ intersects with } C_t\}
\end{aligned}$$

Then, we have $n_1(t) \leq n(t) \leq n_2(t)$.

Also, using area-related arguments, we get $n_1(t) \leq \frac{\pi t^2}{A}$ and $n_2(t) \geq \frac{\pi t^2}{A}$.

Now, consider the diagonal δ of P . If we increase t to $t + \delta$, then all the points counted in $n(t)$ will get counted in $n_1(t + \delta)$. Therefore,

$$n(t) \leq n_1(t + \delta) \leq \frac{\pi(t + \delta)^2}{A} \quad (1)$$

Similarly, if we decrease t to $t - \delta$, then if λ is such that P_λ intersects $C_{t-\delta}$, then such a λ must lie in C_t . Therefore,

$$\frac{\pi(t - \delta)^2}{A} \leq n_2(t - \delta) \leq n(t) \quad (2)$$

Thus, we have from (1) and (2):

$$\frac{\pi(t - \delta)^2}{A} \leq n(t) \leq \frac{\pi(t + \delta)^2}{A}$$

Thus, $n(t) = \frac{\pi t^2}{A} + \mathcal{O}(t)$ which implies the statement of the lemma. \square

1.2 Ideal Class Groups

Definition 1.9. Given a square free positive integer n , we consider the number field:

$$\mathbb{Q}(\sqrt{-n}) = \{a + b\sqrt{-n} : a, b \in \mathbb{Q}\}$$

We denote by \mathcal{O}_{-n} , the ring of integers of $\mathbb{Q}(\sqrt{-n})$, known as the ring of algebraic integers. A simple calculation gives:

$$\mathcal{O}_{-n} = \{a + b\omega_{-n} : a, b \in \mathbb{Z}\}$$

where $\omega_{-n} = \sqrt{-n}$ when $n \equiv 1, 2 \pmod{4}$ and $\omega_{-n} = \frac{1+\sqrt{-n}}{2}$ otherwise.

Instead of working with quadratic number fields, we could work with general algebraic number fields. If F is any such finite extension of \mathbb{Q} and D is its ring of integers, then we have the following result that can be found in a standard algebraic number theory text.

Theorem 1.10. Any ideal I of D consists of a basis of F over \mathbb{Q} . Also, for any ideal I , the quotient ring D/I is finite.

Corollary 1.11. D is a Noetherian ring, i.e. every ascending chain $A_1 \subseteq A_2 \subseteq A_3 \dots$ of ideals terminates.

Proof. Since D/A_1 is finite, there are only finitely many ideals containing A_1 , thus only finitely many distinct ideals in any ascending chain. \square

Corollary 1.12. Every prime ideal of D is maximal.

Proof. Let P be a prime ideal. Then, D/P is a finite integral domain, which in turn, is a field. Hence, P must be a maximal ideal. \square

In particular, we have that \mathcal{O}_{-n} is a Noetherian ring where prime ideals are maximal. This, along with that fact that the ring is also integrally closed, implies that \mathcal{O}_{-n} is a Dedekind domain. We therefore have a unique factorisation of ideals in \mathcal{O}_{-n} and we note this as a theorem.

Theorem 1.13. Given an ideal I of \mathcal{O}_{-n} , we have prime ideals I_1, I_2, \dots, I_r such that $I = I_1 I_2 \dots I_r$ where I_1, I_2, \dots, I_r are unique upto reordering.

We note one final result before we move on to correlate ideals and complex lattices.

Theorem 1.14. For an odd prime p , we have in \mathcal{O}_{-n} :

$$\left(\begin{array}{c} (p) \\ (p) \\ (p) \end{array} \right) = \left\{ \begin{array}{l} (p) \quad \left(\frac{-n}{p} \right) = -1 \\ (p, a + \sqrt{-n})(p, a - \sqrt{-n}) \quad \left(\frac{-n}{p} \right) = 1, a \in \mathbb{Z}, a^2 \equiv -n \pmod{p} \\ (p, \sqrt{-n})^2 \quad \left(\frac{-n}{p} \right) = 0 \end{array} \right\}$$

Also, for 2, we have:

$$(2) = \left\{ \begin{array}{l} (2) \quad n \equiv 3 \pmod{8} \\ (2, \omega_{-n})(2, 1 + \omega_{-n}) \quad n \equiv 7 \pmod{8} \\ (2, \sqrt{-n})^2 \quad n \equiv 2 \pmod{4} \\ (2, 1 + \sqrt{-n})^2 \quad n \equiv 1 \pmod{4} \end{array} \right\}.$$

It is to be noted that all the ideals appearing on the right are prime ideals.

In order to define a group structure on the ideals of \mathcal{O}_{-n} , we define an equivalence relation \sim on the ideals:

$$I \sim J \text{ iff } (a).I = (b).J \text{ for some } a, b, \in \mathcal{O}_{-n}$$

Through this equivalence, the set of all ideals of \mathcal{O}_{-n} gets partitioned into equivalence classes called ideal classes. For given ideals I and J of \mathcal{O}_{-n} , consider their respective ideal classes C_I and C_J . We define an operator $*$ on these as : $C_I * C_J = C_{IJ}$.

It is a simple exercise to check that this operation is well defined and in fact, defines an Abelian group structure on the ideal classes where $C_{\mathcal{O}_{-n}}$, the class of principal ideals, behaves as the identity and the inverse of C_I is given by $C_{I^{-1}}$.

Now, we have all the tools required to see the connection between complex lattices and the ideal class group. We proceed through a series of lemmas whose proofs are a matter of routine algebraic verifications.

Lemma 1.15. Let I be an ideal of \mathcal{O}_{-n} . Then, as a subset of \mathbb{C} , I can be seen as a complex lattice with CM by ω_{-n} . If m is the least positive integer in I and $a + b\sqrt{-n} \in I$ is an element with minimal positive coefficient for $\sqrt{-n}$, then $I = \langle m, a + b\sqrt{-n} \rangle$

Lemma 1.16. Ideals I and J of \mathcal{O}_{-n} are similar over \sim if and only if they are homothetic as lattices, which happens if and only if they have the same

j-invariants.

Lemma 1.17. *Let $\frac{a+\sqrt{-n}}{b}$ be the *j*-invariant of a lattice having CM by ω_{-n} with notation as in Theorems 1.6 and 1.7. Then, the corresponding ideal class having this *j*-invariant as a lattice is the one which contains the ideal $(b, a + \sqrt{-n})$.*

Corollary 1.18. *There is a bijection between classes of homothetic lattices having CM by ω_{-n} and the ideal classes of ideals of \mathcal{O}_{-n} .*

This, without any ambiguity, we can use $\mathcal{Cl}(-n)$ to denote either of the above set of classes, and we refer to the structure obtained as the ideal class group.

Before ending this section, we are going to prove some results about ideals having a given norm. Having fixed a ring of integers \mathcal{O}_{-n} , we denote by x_n the number of its ideals having norm n . It is easy to see that this is a completely multiplicative sequence. By Theorem 1.14, we have for a prime p :

$$x_p = \left\{ \begin{array}{ll} 0 & \left(\frac{-n}{p}\right) = -1 \\ 1 & \left(\frac{-n}{p}\right) = 0 \\ 2 & \left(\frac{-n}{p}\right) = 1 \end{array} \right\}.$$

Now, we try to get some bounds related to the x_i 's. Define for all M , $A_M := \sum_{m=1}^M x_m$. Let w denote the number of units in \mathcal{O}_{-n} . It is an elementary result that $w = 2$ when $n \neq 1, 3$ in which cases, $w = 4, 6$ respectively. For an ideal class \mathcal{C} , define $x_m(\mathcal{C})$ to be the number of ideals of (\mathcal{C}) having norm m .

Let \mathcal{C}_1 denote the class of principal ideals. Consider 2 associates α and α' in \mathcal{O}_{-n} . Then, both of these generate the same principal ideal. Hence, if b_m denotes the number of elements in \mathcal{O}_{-n} having norm m , we have $x_m(\mathcal{C}_1) = \frac{b_m}{w}$.

Now, in order to estimate B_m , the number of elements in \mathcal{O}_{-n} having norm $\leq m$, we make use of Lemma 1.8. Therefore, we have:

$$\left| B_M - \frac{\pi M}{A} \right| \leq C \sqrt{M}$$

where A denotes the area of the parallelogram with sides 1 and ω_{-n} . For $n \equiv 3 \pmod{4}$, $A = \frac{\sqrt{n}}{2}$ and $A = \sqrt{n}$ otherwise. If we use the notation

$A_M(\mathcal{C}) = \sum_{m=1}^M x_m(\mathcal{C})$, we have:

$$|A_M(\mathcal{C}_1) - \frac{\pi M}{Aw}| \leq C' \cdot \sqrt{M}$$

where $C' = C/w$. In fact, we can generalise this in a similar manner for all ideal classes \mathcal{C} and sum this up for all ideals to get a theorem that will be the main ingredient of our final proof:

Theorem 1.19.

$$|A_M - \frac{\pi h(-n)M}{Aw}| \leq C \cdot \sqrt{M}$$

for a suitable constant C for all $M \geq 1$.

1.3 Dirichlet Class Number Formula

In this section, we are going to use the tools developed in the previous sections to prove the celebrated class number formula:

$$\sum_{m=1}^{\infty} \frac{1}{m} \left(\frac{-n}{m} \right) = \frac{h(-n)\pi}{2\sqrt{n}}$$

when $n \equiv 1, 2 \pmod{4}$, and,

$$\sum_{m=1}^{\infty} \frac{1}{m} \left(\frac{-n}{m} \right) = \frac{h(-n)\pi}{\sqrt{n}}$$

when $n \equiv 3 \pmod{4}$, given $n \neq 1, 3$ where, $\left(\frac{-n}{m} \right)$ denotes the generalised Legendre symbol. For $n = 1, 3$, we have the numerator multiplied by 2, 3 respectively in the above formulae.

Before starting with the main course, as an appetizer, we have 2 convergence results from analysis. To recall, a series of the form

$$\sum_{m=1}^{\infty} a_m m^{-s}$$

where the a'_i 's are real numbers, is known as a real Dirichlet series.

Theorem 1.20. *If a_1, a_2, a_3, \dots are real numbers such that there exist $c, r > 0$ such that $\left| \sum_{m=1}^M a_m \right| \leq c.M^r$ for all M , then we have that the Dirichlet series:*

$$\sum_{m=1}^{\infty} a_m m^{-s}$$

converges for all $s > r$ and results in a continuous function in s .

Theorem 1.21. *Let a_1, a_2, a_3, \dots be a completely multiplicative sequence, such that there is a $c > 0$ such that $\sum_{m=1}^M |a_m| \leq c.M$ for all M and $a_p \leq p$ for all primes p , then we have for all $s > 1$:*

$$\sum_{m=1}^{\infty} a_m m^{-s} = \prod_p (1 - a_p p^{-s})^{-1}$$

Now, in order to move towards our required result, we introduce the L -function $L_{-n}(s)$, defined as:

$$L_{-n}(s) = \sum_{m=1}^{\infty} \left(\frac{-n}{m}\right) m^{-s}$$

To see the convergence of this function, we use Theorem 1.21, for which we use the following lemma:

Lemma 1.22. *For any $b \geq 1$, we have:*

$$\sum_b^{b+4n-1} \left(\frac{-n}{m}\right) = 0$$

Proof. As the Legendre symbol is periodic in the top argument with a period $4n$, we assume without loss of generality that $b = 0$. Let the sum on the left hand side be denoted by S . Also, let $k \in (\mathbb{Z}/4n)^*$ be such that $\left(\frac{-n}{k}\right) = -1$. Then, we have

$$\begin{aligned} -S &= \left(\frac{-n}{k}\right) S \\ -S &= \sum_{m \in (\mathbb{Z}/4n)^*} \left(\frac{-n}{m.k}\right) \end{aligned}$$

Now, as m runs through $(\mathbb{Z}/4n)^*$, so does $m.k$. Hence, the sum on the right hand side is nothing but S . Thus, we have the required result, $S = 0$. \square

Hence, we can use Theorem 1.21 for $L_{-n}(s)$ as the above lemma implies that $\sum_1^M \left(\frac{-n}{m}\right) \leq 4n$ for all M . Therefore, we can express the function as:

$$L_{-n}(s) = \prod_p \left(1 - \left(\frac{-n}{p}\right) p^{-s}\right)^{-1}$$

Moving on ahead with our quest, we define the Dedekind Zeta Function of \mathcal{O}_{-n} as the Dirichlet series:

$$\zeta_{-n}(s) = \sum_{m=1}^{\infty} x_m m^{-s}$$

where the x_i 's are as defined in the previous section. Then, we have the following proposition which will put us on track for our final result:

Theorem 1.23. *The Dedekind Zeta Function converges for all $s > 1$ and can be expressed as:*

$$\zeta_{-n}(s) = \prod_{p, \left(\frac{-n}{p}\right)=1} (1 - p^{-s})^{-2} \cdot \prod_{p, \left(\frac{-n}{p}\right)=0} (1 - p^{-s})^{-1} \cdot \prod_{p, \left(\frac{-n}{p}\right)=-1} (1 - p^{-2s})^{-1} \quad \square$$

Also, $\lim_{s \rightarrow 1^+} (s - 1)\zeta_{-n}(s) = \frac{h(-n)\pi}{Aw} \quad \square$

Proof. The convergence follows directly from the results of the previous section and Theorems 1.20 and 1.21. In order to calculate the limit, we define:

$$f(s) = \sum_{m=1}^{\infty} \left(x_m - \frac{h(-n)\pi}{Aw}\right) m^{-s}$$

Then, using the results from Theorems 1.19 and 1.20, we have that $f(s)$ converges for all $s > \frac{1}{2}$. Furthermore, for $s > 1$, we have:

$$\zeta_{-n}(s) = f(s) + \frac{h(-n)\pi}{Aw} \zeta(s)$$

where $\zeta(s)$ denotes the Riemann Zeta Function. Therefore, we have:

$$\lim_{s \rightarrow 1^+} (s - 1)\zeta_{-n}(s) = \lim_{s \rightarrow 1^+} (s - 1)f(s) + \lim_{s \rightarrow 1^+} (s - 1) \frac{h(-n)\pi}{Aw} \zeta(s)$$

and, as f is continuous at $s = 1$ and $\lim_{s \rightarrow 1^+} (s - 1)\zeta(s) = 1$, we have what we set out to prove. \square

Theorem 1.24. *For $s > 1$,*

$$\zeta_{-n}(s) = \zeta(s)L_{-n}(s)$$

The proof of the above theorem is an easy exercise of comparing the Euler Products for both sides which have already been determined. And so, the wait ends as we have our final result.

Theorem 1.25. *(Dirichlet Class Number Formula)*

$$L_{-n}(1) = \frac{h(-n)\pi}{Aw}$$

Proof.

$$L_{-n}(1) = \lim_{s \rightarrow 1^+} L_{-n}(s)$$
$$L_{-n}(1) = \lim_{s \rightarrow 1^+} \frac{(s-1)\zeta_{-n}(s)}{(s-1)\zeta(s)}$$

Thus, Theorem 1.23 puts the final nail in the coffin and the theorem stands proven. □



2 Analytic Number Theory

In this section, we move on to analytic number theory. The first subsection deals with Dirichlet's theorem on the infinitude of primes in arithmetic progressions. The second one talks about Brun's theorem on the convergence of the sum of reciprocals of twin primes.

2.1 Dirichlet's Theorem on the infinitude of primes in arithmetic progressions

2.1.1 Dirichlet Characters

Definition 2.1. *Given a finite Abelian group G , the characters of G are the group homomorphisms χ from G to \mathbb{C}^**

For any G , we always have the trivial homomorphism χ_0 known as the principal character defined as:


$$\chi_0(g) = 1 \quad \forall g \in G$$

Now, suppose we have a cyclic group G with generator g_0 having order n and let χ be character for G . Then, we must have:

$$\chi(g_0)^n = \chi(g_0^n) = \chi(1) = 1$$

And thus, $\chi(g_0)$ must be an n^{th} root of unity (not necessarily primitive). Also, given ω , any n^{th} of unity, we get a character χ of G by putting $\chi(g_0) = \omega$. This classifies all characters for cyclic groups. In fact, if an Abelian group can be written as a direct product of cyclic groups, we can classify all the characters in a similar manner.

But then, the Fundamental Theorem of Finite Abelian Groups says exactly that every group is a direct product of cyclic groups (of prime power orders). Thus, we have all characters classified for finite Abelian groups.

We can now associate a dual group of a group G , denoted by \hat{G} with the help of these characters with inverse and identity being defined naturally. Next we state 2  orthogonality results that relate the characters of a given group.

Lemma 2.2. If χ and ψ are characters of a group G , we have

$$\sum_{g \in G} \chi^{-1}(g)\psi(g) = \begin{cases} |G| & \text{if } \chi = \psi \\ 0 & \text{otherwise} \end{cases}.$$

Lemma 2.3. If g and h are elements of a group G , we have

$$\sum_{\chi \in \hat{G}} \chi(g)\chi(h^{-1}) = \begin{cases} |G| & \text{if } h = g \\ 0 & \text{otherwise} \end{cases}.$$

Next, we move on to define Dirichlet characters. If in the above discussion, take $G = \mathbb{Z}/n\mathbb{Z}^*$. Consider any $\chi \in \hat{G}$, then we define a function $\tilde{\chi}$ over \mathbb{Z} , such that for $a \in \mathbb{Z}$, if $(a, n) = 1$, we put $\tilde{\chi}(a) = \chi(a \pmod n)$. If $(a, n) > 1$, we put $\tilde{\chi}(a) = 0$. Now, it is elementary to see that this $\tilde{\chi}$ (which we will be referring to as χ by abuse of notation) is periodic and completely multiplicative.

These functions defined above are known as Dirichlet characters and will play a fundamental role in the proof.

2.1.2 L-series

Given a Dirichlet character $\chi \pmod q$, we associate with it its L -series:

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

By the following result, we have the convergence of $L(s, \chi)$ for non-negative s . (specifically at $s = 1$)

Lemma 2.4. Let χ be a non-trivial Dirichlet character mod q , then the L -series converges for $s > 0$. In fact, we have:

$$\sum_{n > x} \frac{\chi(n)}{n^s} = \mathcal{O}\left(\frac{1}{n^s}\right)$$

Before moving further, we recall the Von Mangoldt function Λ defined to be equal to $\log p$ for powers of a prime p and 0 otherwise. It is trivial to see that for any n , we have $\log n = \sum_{d|n} \Lambda(d)$.

Next, we are going to discuss the behavior of the sum $\sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n}$. Consider a non-trivial Dirichlet character χ such that $L(1, \chi) \neq 0$. We have:

$$\sum_{n \leq x} \frac{\chi(n) \log n}{n} = \sum_{n \leq x} \frac{\chi(n)}{n} \sum_{d|n} \Lambda(d)$$

By the multiplicativity of Λ , we have

$$\sum_{n \leq x} \frac{\chi(n) \log n}{n} = \sum_{d \leq x} \frac{\chi(d)\Lambda(d)}{d} \sum_{e \leq x/d} \frac{\chi(e)}{e}$$

The inner sum is $L(1, \chi) - \sum_{e > x/d} \frac{\chi(e)}{e}$ which is $L(1, \chi) + \mathcal{O}(d/x)$ by Lemma 2.4. Thus:

$$\sum_{n \leq x} \frac{\chi(n) \log n}{n} = L(1, \chi) \sum_{d \leq x} \frac{\chi(d)\Lambda(d)}{d} + \mathcal{O}\left(\frac{1}{x} \sum_{d \leq x} \Lambda(d)\right)$$

By a well known result, $\sum_{d \leq x} \Lambda(d) = \mathcal{O}(x)$ and so, the above error becomes $\mathcal{O}(1)$. Now, by using Euler's partial summation technique, it is not difficult to see that the left hand side of the above equation is $\mathcal{O}(1)$. Thus, as $L(1, \chi) \neq 0$, we have our result:

$$\sum_{d \leq x} \frac{\chi(d)\Lambda(d)}{d} = \mathcal{O}(1)$$

Now, suppose χ is a non-trivial Dirichlet character such that $L(1, \chi) = 0$. Now, let's try to estimate the sum. First, by Mobius inversion, we have,

$$\Lambda(n) = - \sum_{d|n} \mu(d) \log d$$

$$\Lambda(n) + \log x \sum_{d|n} \mu(d) = \sum_{d|n} \mu(d) \log(x/d)$$

Using this fact, we move towards our required sum:

$$\log x + \sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = \sum_{n \leq x} \frac{\chi(n)}{n} \sum_{d|n} \mu(d) \log(x/d)$$

As before, using multiplicativity of χ , we have:

$$\log x + \sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = \sum_{d \leq x} \mu(d) \log \frac{x \chi(d)}{d} \sum_{e \leq x/d} \frac{\chi(e)}{e}$$

$$\log x + \sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n} = L(1, \chi) \sum_{d \leq x} \mu(d) \log \frac{x \chi(d)}{d} + \mathcal{O}(1)$$

where the error term is $\mathcal{O}(1)$ with the help of Stirling's approximation. Thus, as $L(1, \chi) = 0$, we have:

$$\sum_{d \leq x} \frac{\chi(d)\Lambda(d)}{d} = -\log x + \mathcal{O}(1)$$

Finally, in order to estimate sum for the trivial character χ_0 , we use Chebyshev's following approximation:

$$\sum_{d \leq x} \frac{\Lambda(d)}{d} = \sum_{p^k \leq x} \frac{\log p}{p^k}$$

Using this estimate, we have

$$\sum_{d \leq x} \frac{\chi_0(d)\Lambda(d)}{d} = \log x + \mathcal{O}(1)$$

Now, we are ready to establish results about non-vanishing of $L(1, \chi)$. We refer to a character χ that assumes at least one non-real value as a complex character. Then we have the following result.

Lemma 2.5. *For a complex character $\chi \pmod{q}$, $L(1, \chi) \neq 0$.*

Proof.

$$\sum_{\chi} \sum_{n \leq x} \frac{\Lambda(n)\chi(n)}{n} = (1 - k) \log x + \mathcal{O}(1)$$

where k denotes the number of characters χ such that $L(1, \chi) = 0$. Now, if we interchange the order of the summation and use the orthogonality relations, we have:

$$\sum_{\chi} \sum_{n \leq x} \frac{\Lambda(n)\chi(n)}{n} = \frac{1}{\phi(q)} \sum_{n \leq x, n \equiv 1 \pmod{q}} \frac{\Lambda(n)}{n} \geq 0$$

Thus, we must have $k \leq 1$. Now, for a complex character χ , if $L(1, \chi) = 0$, then it is easy to see that we will also be able to say that for its complex conjugate $\bar{\chi}$. But, as $k \leq 1$, we can not have $L(1, \chi) = 0$ and the result stands proven. \square

In order to establish the non-vanishing of $L(1, \chi)$ for real characters χ , we'll need some more work. We define for $0 < x < 1$:

$$f(x) = \sum_{d=1}^{\infty} \sum_{k=1}^{\infty} \chi(d)x^{kd}$$

By the convergence of the geometric series and using the fact that $|\chi(d)| \leq 1$, we have that the above sum converges absolutely. As a result, we can reorder the terms in the sum and write this as:

$$f(x) = \sum_{d=1}^{\infty} \chi(d) \frac{x^d}{1-x^d} = \sum_{n=1}^{\infty} c_n x^n$$

where

$$c_n = \sum_{d|n} \chi(d)$$

Now, as $\chi(n)$ is a multiplicative function, so is c_n and so, we can show that each of the c_i 's are non-negative if we can check that just for prime powers. Now, as χ is a real character, the only values it takes are $-1, 0$ and 1 and so, it is an easy exercise to see that the sum $1 + \chi(p) + \chi(p^2) + \dots + \chi(p^e)$ is always non-negative.

In fact, whenever e is even above, the sum is ≥ 1 and so, whenever n is a perfect square, we have $c_n \geq 1$. Now, as the number of perfect squares is unbounded, there are too many coefficients in $f(x)$ that become greater than or equal to 1 and so, $\lim_{x \rightarrow 1^-} f(x) = \infty$.

Now, suppose $L(1, \chi) = 0$. Then, in order to derive a contradiction, we observe that we can write $-f(x) = \frac{L(1, \chi)}{1-x} - f(x)$ which can be expressed as:

$$-f(x) = \sum_{n=1}^{\infty} \left(\frac{1}{n(1-x)} - \frac{x^d}{1-x^d} \right) = \sum_{n=1}^{\infty} b_n(x) \chi(n)$$

By an application of the arithmetic mean - geometric mean inequality and induction, one can see that $b_n(x)$ forms a non-increasing sequence of functions, which goes to zero as $n \rightarrow \infty$. Hence, we have that the sequence consists of non-negative functions only. Next,

$$\sum_{n=1}^M b_n(x) \chi(n) = S(M) b_m(x) + \sum_{n=1}^{M-1} S(n) (b_n(x) - b_{n+1}(x))$$

where $S(n) = \sum_{k=1}^n \chi(k)$. Using the fact that S is bounded above by q , we have for the above sum:

$$\left| \sum_{n=1}^M b_n(x) \chi(n) \right| \leq q b_1(x) = q \left(\frac{1}{1-x} - \frac{x}{1-x} \right) = q$$

Thus, we have that the partial sums of the series $\sum_{n=1}^{\infty} b_n(x) \chi(n)$ are bounded above by q and therefore, so must be the entire sum. Thus, we have $|f(x)| \leq 1$. But then, we had proved that f becomes unbounded near 1. Therefore, we have a contradiction and so $L(1, \chi) \neq 0$.

Thus, combined with our previous lemma, we have that for any non-trivial character χ , we have $L(1, \chi) \neq 0$. This is all we need to attack at our theorem with full force.

2.1.3 The theorem

By now, we have virtually proven what we set out for and we only need to clear the fog from our spectacles to get the final view. We have proven that for a non-trivial Dirichlet character mod q , we have

$$\sum_{d \leq x} \frac{\chi(d)\Lambda(d)}{d} = \mathcal{O}(1)$$

and for the trivial character χ_0 , we have

$$\sum_{d \leq x} \frac{\chi_0(d)\Lambda(d)}{d} = \log x + \mathcal{O}(1)$$

Thus, we have for a such that $(a, q) = 1$:

$$\sum_{n \leq x, n \equiv a \pmod{q}} \frac{\Lambda(n)}{n} = \frac{1}{\phi(q)} \sum_{\chi} \chi^{-1}(a) \sum_{n \leq x} \frac{\chi(n)\Lambda(n)}{n}$$

Using the above results, we get:

$$\sum_{n \leq x, n \equiv a \pmod{q}} \frac{\Lambda(n)}{n} = \frac{1}{\phi(q)} \log x + \mathcal{O}(1)$$

But, by Chebyshev's approximations, it is not hard to show that:

$$\sum_{n \leq x, n \equiv a \pmod{q}} \frac{\Lambda(n)}{n} = \sum_{p \leq x, p \equiv a \pmod{q}} \frac{\log p}{p} + \mathcal{O}(1)$$

And thus, we have our grand result

Theorem 2.6. (*Dirichlet's Theorem*) Given any q and a such that $(a, q) = 1$, we have

$$\sum_{p \leq x, p \equiv a \pmod{q}} \frac{\log p}{p} = \frac{1}{\phi(q)} \log x + \mathcal{O}(1)$$

where the sum is taken over primes. In particular, there are infinitely many primes of the form $qn + a$.



2.2 Brun's Theorem on the convergence of the sum of reciprocals of twin primes

2.2.1 Brun's Simple Pure Sieve

We start off with some notations that we will be using throughout. By \mathcal{A} , we denote a finite sequence of positive integers $\{a_i\}$. Let X denote an approximate size of this set. Let \mathcal{P} denote a set of primes. Then we define:

$$S(\mathcal{A}, \mathcal{P}, z) = |\{a \in \mathcal{A} : (a, P(z)) = 1\}|$$

where $P(z) = \prod_{p \in \mathcal{P}, p \leq z} p$. If \mathcal{P} is finite, we can strike off z from the above notation and talk about $S(\mathcal{A}, \mathcal{P})$ where the product is taken over all primes in \mathcal{P} . Next we use A_d to denote the number of elements in \mathcal{A} that are divisible by d . Now, in order to derive some concrete results about $S(\mathcal{A}, \mathcal{P}, z)$, we assume the existence of a multiplicative function α taking values in $[0, 1]$ and a function r such that for all d ,

$$A_d = X\alpha(d) + r(d)$$

Next, let us look at a combinatorial result before we move on to the sieve.

Lemma 2.7. *Let a_1, a_2, \dots, a_n be a sequence of real numbers in $[0, 1]$. Then we have,*

$$\sum_{k=1}^m (-1)^k \sigma_k(a_1, a_2, \dots, a_n) - \prod_{j=1}^n (1 - a_j)$$

is nonnegative or nonpositive accordingly as m is even or odd. Here, σ_k denotes the k^{th} symmetric function on the n variables, which becomes 0 for $k > n$.

The proof of this lemma is an easy exercise in induction. A simple corollary of this lemma which we will be needing is:

Corollary 2.8. *Let X be a non-empty set with n elements. Let P_1, P_2, \dots, P_r be properties these elements may have. For any $I \subseteq \{1, 2, \dots, r\}$, denote by $N(I)$ the number of elements having the properties having their indices in I . If N_0 denotes the number of elements having none of the properties, then*

we have:

$$N_0 \leq \sum_{k=0}^m (-1)^k \sum_{I \subseteq \{1,2,\dots,r\}, |I|=k} N(I)$$

for m even. For odd m , we have the reverse inequality:

$$N_0 \geq \sum_{k=0}^m (-1)^k \sum_{I \subseteq \{1,2,\dots,r\}, |I|=k} N(I)$$

A direct consequence of this corollary is our main result of this section which states some bounds on $S(\mathcal{A}, \mathcal{P})$. Using the aforementioned notation, we have the Brun's simple pure sieve:

$$\sum_{d|P, \nu(d) \leq m-1} \mu(d) A_d \leq S(\mathcal{A}, \mathcal{P}) \leq \sum_{d|P, \nu(d) \leq m} \mu(d) A_d$$

where m is an even integer and μ denotes the Mobius function. In order to derive some more insight, we try to quantify the above found in the following theorem.

Theorem 2.9. *For all even integers $m \geq 0$,*

$$S(\mathcal{A}, \mathcal{P}) = X \prod_{p \in \mathcal{P}} (1 - \alpha(p)) + \mathcal{O}\left(\sum_{d|P, \nu(d) \leq m} |r(d)|\right) + \mathcal{O}\left(X \sum_{d|P, \nu(d) \geq m} \alpha(d)\right)$$

Proof. By the above inequalities, we have:

$$S(\mathcal{A}, \mathcal{P}) = \sum_{d|P, \nu(d) \leq m} \mu(d) A_d + \mathcal{O}\left(\sum_{d|P, \nu(d) = m} A_d\right)$$

Using the expansion $A_d = X\alpha(d) + r(d)$, we have:

$$S(\mathcal{A}, \mathcal{P}) = X \sum_{d|P, \nu(d) \leq m} \mu(d) \alpha(d) + \mathcal{O}\left(\sum_{d|P, \nu(d) \leq m} |r(d)|\right) + \mathcal{O}\left(X \sum_{d|P, \nu(d) = m} \alpha(d)\right)$$

Now, in order to turn the main term into a product, we need to add the corresponding terms in the sum where $\nu(d) > m$. This introduces an error of at most $\mathcal{O}\left(X \sum_{d|P, \nu(d) > m} \alpha(d)\right)$. Adding this to the final error term, we get our

final result:

$$S(\mathcal{A}, \mathcal{P}) = X \prod_{p \in \mathcal{P}} (1 - \alpha(p)) + \mathcal{O}\left(\sum_{d|P, \nu(d) \leq m} |r(d)|\right) + \mathcal{O}\left(X \sum_{d|P, \nu(d) \geq m} \alpha(d)\right)$$

which was what we wanted. □

Before ending, we state here a weakened form of a result proved by Merten, which, although not related to Brun's sieve theory, will be used by us in the proof of our main theorem. This theorem is a corollary of the result that:

$$\sum_{p \leq x} \frac{1}{p} = \log \log x + C + \mathcal{O}\left(\frac{1}{\log x}\right)$$

(for a constant C which won't be relevant in our discussion.)

Theorem 2.10. *For all $x \geq 2$, we have*

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) = \theta\left(\frac{1}{\log x}\right)$$

Proof. Let $P_x = \prod_{p \leq x} \left(1 - \frac{1}{p}\right)$. Then,

$$\log P_x = \sum_{p \leq x} \log \left(1 - \frac{1}{p}\right) = - \sum_{p \leq x} \frac{1}{p} - \sum_{p \leq x} \sum_{k=2}^{\infty} \frac{1}{kp^k}$$

where we have used the Taylor series for $\log(1+x)$. It is easy to see that the 2^{nd} sum above converges (to say S), and so, managing the error terms, we have

$$\log P_x = - \log \log x - C - S + \mathcal{O}\left(\frac{1}{\log x}\right)$$

Thus, taking antilogs, we have our required result. □

2.2.2 The theorem

In this section, we prove results using the theory developed in the previous section, which will lead us eventually to Brun's theorem. Before that, we introduce some more notation. By $\pi_2(z)$, we denote the number of twin prime pairs with at least one prime below z . Next we define its generalisation:

$$\pi_2(x, z) = |\{n \leq x : p|n(n+2) \Rightarrow p > z\}|$$

Our strategy is going to be as follows. We first derive a bound on $\pi_2(x, z)$. This will then determine a bound on $\pi_2(z)$. The final bound will lead us

to our convergence result. So, we start with the following theorem which is going to provide the main blow towards our theorem.

Theorem 2.11. *Suppose z as a function of x goes to ∞ as x goes to ∞ and at the same time, $z(x) \leq x^{1/20 \log \log x}$ for large x . Then, we have:*

$$\pi_2(x, z) = \theta(x / \log^2 z)$$

Proof. Using the notation as in the previous section, if we take $\mathcal{A} = \{a_n = n(n+2) : n \leq x\}$ and \mathcal{P} to be the set of all primes, then $S(\mathcal{A}, \mathcal{P}, z)$ becomes equal to $\pi_2(x, z)$. Thus, X can be taken to be x and $\alpha(d)$ becomes equal to $\omega(d)/d$ where $\omega(d)$ denotes the number of roots of the polynomial $n(n+2) = 0$ in $\mathbb{Z}/d\mathbb{Z}$. (Thus, $\omega(d) \leq 2$). The multiplicativity of ω implies the multiplicativity of α .

Next, we put $r(d) = A_d - x\alpha(d)$. It is trivial to see that $r(d) \leq |\omega(d)| = \prod_{p|d} \omega(p) \leq 2^{\nu(d)}$. Now, we are ready to make use of Brun's sieve. By Theorem 2.9, we have:

$$\pi_2(x, z) = x \prod_{p \leq z} (1 - \alpha(p)) + \mathcal{O}\left(\sum_{d|P, \nu(d) \leq m} 2^{\nu(d)}\right) + \mathcal{O}\left(x \sum_{d|P, \nu(d) \geq m} \alpha(d)\right)$$

for any even integer m . Set $m = 10 \lfloor \log \log z \rfloor$. Now, we try to estimate each term in the above sum separately.

For any prime p , the given quadratic has exactly 2 solutions, except $p = 2$, when it has only 1 solution. Thus,

$$x \prod_{p \leq z} (1 - \alpha(p)) = x/2 \prod_{p \leq z} (1 - 2/p)$$

Now, making use of Theorem 2.10, we can write this expression as:

$$x \prod_{p \leq z} (1 - \alpha(p)) = 2x \prod_{2 < p \leq z} \frac{1 - 2/p}{(1 - 1/p)^2} \prod_{p \leq z} (1 - 1/p)^2 = \theta(x / \log^2 z)$$

where we use the fact that the product $\prod_{2 < p \leq z} \frac{1 - 2/p}{(1 - 1/p)^2}$ converges to a non-zero constant.

Next, let us estimate the error terms. We have first $\mathcal{O}\left(\sum_{d|P, \nu(d) \leq m} 2^{\nu(d)}\right)$. Now, any d for which $\nu(d) = k$ for a given k is given by choosing k primes less than or equal to z . This number is given by $\binom{\pi(z)}{k}$. Thus, using the inequality $\binom{n}{k} \leq n^k$ the given sum becomes:

$$\mathcal{O}\left(\sum_{d|P, \nu(d) \leq m} 2^{\nu(d)}\right) = \sum_{k=0}^m 2^k \binom{\pi(z)}{k} \leq \sum_{k=0}^m (2\pi(z))^k \leq 2(2\pi(z))^m \leq 2z^m$$

Now, we had taken $m = 10 \lfloor \log \log z \rfloor$ and so, the given sum become less than or equal to:

$$2z^{10 \log \log z} \leq 2z^{10 \log \log x} \leq 2\sqrt{x}$$

by what was initially given. Now, as $z \leq x$, we have:

$$2\sqrt{x} = o\left(\frac{x}{\log^2 x}\right) = o\left(\frac{x}{\log^2 z}\right)$$

Therefore, this error term is $o(x/\log^2 z)$ and thus, goes to zero when compared to the main term. Now, we move on to the final error term $x \sum_{d|P, \nu(d) \geq m} \alpha(d)$.

Now, for a given k , we have

$$\sum_{d|P, \nu(d)=k} \alpha(d) = \sum_{p_1 < p_2 < \dots < p_k \leq z} \prod_{i=1}^k \alpha(p_i) \leq \frac{1}{k!} \left(\sum_{p \leq z} \alpha(p)\right)^k$$

because each term on the LHS occurs $k!$ times in the multinomial expansion of the RHS. Now, for each p , $\alpha(p) \leq 2/p$. Thus using the estimate $\sum_{p \leq z} \frac{1}{p} \leq$

$\log \log z + c$, we have:

$$\frac{1}{k!} \left(\sum_{p \leq z} \alpha(p)\right)^k \leq \frac{1}{k!} (2 \log \log z + 2c)^k$$

Now, if we take the sum over all k , the ratio of consecutive terms in the sum tends to less than $1/2$ as z , and hence x , increases. Thus, the entire sum is bounded above by twice the first term by approximation as a geometric sum. Thus,

$$\sum_{k \geq m} \frac{1}{k!} \left(\sum_{p \leq z} \alpha(p)\right)^k \leq \frac{2}{m!} (2 \log \log z + 2c)^m \leq 2 \left(\frac{2e \log \log z + 2ec}{m}\right)^m$$

where the 2nd inequality follows from the fact that $e^m > m^m/m!$ which follows from the Taylor expansion of e^m . Now, as $m = 10\lfloor \log \log z \rfloor$, the inner term above is bounded above by a constant above $2e/10$, say $3/5$. Thus, our error term is less than:

$$x(3/5)^m = x(3/5)^{10\lfloor \log \log z \rfloor} = \mathcal{O}(x/\log^5 z)$$

Thus, this error term too will go to zero as compared to the main term. Hence, we have our final result

$$\pi_2(x, z) = \theta(x/\log^2 z)$$

□

Now, we use the above theorem to put a bound on $\pi_2(x)$. For any choice of z , we have

$$\pi_2(x) \leq z + \pi_2(x, z)$$

Taking $z = z(x) = x^{1/20 \log \log x}$ and applying the above theorem, for $x \rightarrow \infty$, we have:

$$\pi_2(x) = \mathcal{O}\left(x^{1/20 \log \log x} + \frac{x}{\log^2 x}(\log \log x)^2\right) = \mathcal{O}\left(\frac{x}{\log^2 x}(\log \log x)^2\right)$$

In fact, for our purposes, we'll only be needing a weaker bound on $\pi_2(x)$ implied by the above bound:

$$\pi_2(x) = \mathcal{O}\left(\frac{x}{\log^{3/2} x}\right)$$

Now, we prove our final theorem.

Theorem 2.12. (*Brun's Theorem*) *If it is an infinite sum,*

$$\sum_p \frac{1}{p}$$

converges, where the sum is taken over all primes p such that $p + 2$ is also a prime.



Proof. Let p_n denote the n^{th} prime such that $p + 2$ is also a prime. Then, by the above bound, we have:

$$n = \pi_2(p_n) \leq \frac{kp_n}{\log^{3/2} p_n}$$

for some sufficiently large constant k . Therefore,

$$kp_n \geq n \log^{3/2} p_n \geq n \log^{3/2} n$$

Now, the sum $\sum_{n=2}^{\infty} \frac{1}{n \log^{3/2} n}$ is known to converge by Cauchy condensation test. Hence, by the above inequality, comparison test does our job. \square

References

- [1] Kenneth Ireland, Michael Rosen, A Classical Introduction to Modern Number Theory, 1990
- [2] M. Ram Murty, Jody Esmonde, Problems in Algebraic Number Theory, 2004
- [3] Tom Weston, Lectures on the Dirichlet Class Number Formula for Imaginary Quadratic Fields, 2004
- [4] David M. Burton, Elementary Number Theory, 2007
- [5] David S. Dummit, Richard M. Foote, Abstract Algebra, 2004
- [6] Paul Pollack, Not Always Buried Deep, Selections from Analytic and Combinatorial Number Theory, 2003