5-2008

# The Square Threshold Problem in Number Fields

Matt Lafferty
*Clemson University*, mlaffer@clemson.edu

# The Square Threshold Problem in Number Fields

A Master's Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Mathematical Sciences

by
Matthew John Lafferty
May 2008

Accepted by:
Dr. Kevin James, Committee Chair
Dr. Neil J. Calkin
Dr. Hiren Maharaj

# Abstract

Let $K$ be a degree $n$ extension of $\mathbb{Q}$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. Let $x \geq 2$. Suppose we were to generate an ideal sequence by choosing ideals from $\{I \subseteq \mathcal{O}_K : N(I) \leq x\}$, independently and with uniform probability. How long would our sequence of ideals need to be before we obtain a subsequence whose terms have a product that is a square ideal in $\mathcal{O}_K$? We show that the answer is about $\exp \sqrt{2 \ln(x) \ln \ln(x)}$.

# Table of Contents

# Chapter 1

# Introduction

## 1.1 The Square Threshold Problem in $\mathbb{Z}$

Suppose we were to construct a sequence by randomly selecting integers independently and with uniform probability from the set $\{1\ldots,x\}$. How many integers do we need to select before the product of the terms of some subsequence is a square? For example, consider the following sequence of integers drawn from the set $\{1,\ldots,9\}$,

$$2,7,5,3,6,3,8,9,9,4,1,7,3,9,3,2,5,\ldots$$

The first 5 terms of this sequence contain a subsequence for which the product of the terms is a square. Namely, the subsequence $2,6,3$, whose product is $36=6^2$. Sequences having the property that the product of their terms is a square will be referred to as *square dependent.*

Given an integer sequence $\mathcal{N}$, constructed by selecting integers independently and with uniform probability from the set $\{1\ldots,x\}$, let $\Omega$ be the smallest positive integer such that the first $\Omega$ terms of $\mathcal{N}$ contain a square dependent subsequence. In the example given above, $\Omega=5$ as the first 5 terms contain a square dependent subsequence, and the first 4 do not. The integer $\Omega$ is said to be the *threshold* in which our sequence of integers goes from not having a square dependent subsequence, to having a square dependent subsequence.

The Square Threshold problem in $\mathbb{Z}$ is concerned with determining probabilistic bounds for $\Omega$ for arbitrary sequences of integers drawn from the set $\{1, \ldots, x\}$ in the manner specified above, where $x \geq 2$. This means that we want to determine an interval $(A(x), B(x))$, dependent on $x$, such that the probability $\Omega \in (A(x), B(x))$ tends to 1 as $x \to \infty$.

While applicability is not a requisite quality for interesting mathematical problems, a primary motivation in determining probabilistic bounds for $\Omega$ has been the heuristic analysis of the running times for integer factorization algorithms. In many of these algorithms, for example the Quadratic Sieve, a sequence of integers is generated until a square dependent subsequence is obtained. As such, an estimate on how long a generated sequence needs to be in order to obtain a square dependent subsequence is essential in estimating the overall running time of the algorithm. While the sequence of integers generated by these integer factorization algorithms may not be random in the sense we are considering, for the purposes of a heuristic analysis it is fruitful to assume so [4].

## 1.2 Current Progress

The first probabilistic bound on $\Omega$ has been attributed to R. Schroeppel [4], who showed for $\epsilon > 0$, the probability that $\Omega < (1-\epsilon)L^{\sqrt{2}}$ tends to 1 as $x \to \infty$, where $L = \exp(\sqrt{\ln(x) \ln\ln(x)})$. While Schroeppel's own proof has not been published, a proof of this upper bound can be found in [11, Prop. 4.1].

In 1994, Pomerance gave a probabilistic lower bound for $\Omega$, showing that for $\epsilon > 0$, the probability $\Omega \leq L^{\sqrt{2}-\epsilon}$ tends to 1 as $x \to \infty$ [13, Thm. 1]. In the same paper, he also conjectured the existence of threshold function $T(x)$ such that for $\epsilon > 0$, the probability that $(1-\epsilon)\,T(x) < \Omega < (1+\epsilon)\,T(x)$ tends to 1 as $x \to \infty$. In this direction, Croot, Granville, Pemantle, and Tetali [4] have recently made the following conjecture,

**Conjecture 1.** *For every $\epsilon > 0$, the probability that*

$$(e^{-\gamma} - \epsilon)L^{\sqrt{2}} \;\leq\; \Omega \;\leq\; (e^{-\gamma} + \epsilon)L^{\sqrt{2}},$$

*tends to 1 as $x \to \infty$, where $\gamma = 0.577\ldots$ is the Euler-Mascheroni constant.*

They prove the upper bound given in Conjecture 1, and a slightly weaker lower bound of $(\pi/4)(e^{-\gamma} - \epsilon)L^{\sqrt{2}}$.

## 1.3 The Square Threshold Problem in $\mathcal{O}_K$

Let $K$ be a degree $n$ extension of $\mathbb{Q}$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. While $\mathcal{O}_K$ is not necessarily a unique factorization domain, it is well known that $\mathcal{O}_K$ is a Dedekind domain. This implies that every ideal $I \subseteq \mathcal{O}_K$ can be written uniquely (up to order) as the product of powers of prime ideals. Because of this, many of the arguments used to determine probabilistic bounds for $\Omega$ with respect to the integers, naturally generalize to ideals of $\mathcal{O}_K$.

Define $S(x) = \{I \subseteq \mathcal{O}_K : N(I) \leq x\}$. Suppose we were to generate an ideal sequence by choosing ideals from $S(x)$ independently and with uniform probability (that is, in the same manner previously described for generating integer sequences). How long would our sequence of ideals need to be before we obtain a square dependent subsequence? Analogous to the integer case, a sequence of ideals of $\mathcal{O}_K$ is said to be square dependent if the product of its terms is the square of an ideal in $\mathcal{O}_K$. Note that this is precisely the question we asked with respect to integers, if we let $K = \mathbb{Q}$.

Our primary goal in this paper is to prove the following theorem,

**Theorem 1.** *Let $K$ be a degree $n$ extension of $\mathbb{Q}$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. Let $\Omega$ be the smallest integer such that the first $\Omega$ terms of an ideal sequence generated by randomly and independently selecting ideals with uniform probability from the set $S(x)$ contains a square dependent subsequence. Then for any $\epsilon > 0$ the probability that $\Omega$ is in the interval*

$$\left(L^{\sqrt{2}-\epsilon},\, L^{\sqrt{2}+\epsilon}\right)$$

*tends to 1 as $x \rightarrow \infty$.*

The proof of Theorem 1 is given in Chapter 3. Central to the proof of Theorem 1 is the concept of a smooth number, and its generalization, the smooth ideal. In Chapter 2 we introduce the concept of smooth numbers and smooth ideals, and results pertaining to these concepts that will be essential to our proof of Theorem 1.

# Chapter 2

# Smooth Numbers & Smooth Ideals

## 2.1 Smooth Numbers

Generally speaking, an integer is said to be smooth if it has no large prime factors [12]. However, this definition is a little vague as to what constitutes a "large" prime factor. We will over come this ambiguity with the following definition,

**Definition.** *For $y \geq 2$, we say that an integer $x$ is $y$-smooth if for every prime $p \mid x$ we have $p \leq y$. In other words, $x$ is said to be $y$-smooth if $x$ has no prime factors exceeding $y$.*

For our purposes, we will be interested in the number of $y$-smooth integers below a bound $x$, for a given $x$ and $y$. We denote the number of such integers by $\Psi(x, y)$. There is a great deal of literature on $\Psi(x, y)$, which speaks in large part to the importance of smooth numbers in computational number theory. For excellent surveys on $\Psi(x, y)$ see [7] and [6]. The earliest, and perhaps most fundamental result pertaining to $\Psi(x, y)$ is due to K. Dickman. In 1930, Dickman showed that the proportion of integers less than $x$ that are $x^{1/u}$ smooth, where $u > 0$, tends to a non zero limit as $x \to \infty$ [6]. Specifically he proved the following theorem,

**Theorem 2.1.1.** *For each real number $u > 0$, there is a number $\rho(u) > 0$ such that*

$$\Psi(x, x^{1/u}) \sim x\rho(u).$$

In keeping with Dickman's notation, we will take $y = x^{1/u}$ whenever the function $\Psi(x, y)$ is considered. In 1949, V. Ramaswami gave a rigorous proof that $\rho(u)$ is in fact a continuous function for $u > 0$ [14]. He proved the following,

**Theorem 2.1.2.** *A function $\rho(u)$ defined for all $u > 0$ exists such that*

1.) *$\rho(u) > 0$ for $u > 0$,*

2.) *$\rho(u)$ is continuous for $u > 0$, and*

3.) *for any fixed $u$, $\Psi(x, x^{1/u}) = x\rho(u) + O(x/\ln(x))$*

The function $\rho$ is commonly referred to as the Dickman-de Bruijn function. In section 2.4 we give a proof of a result similar to Ramaswami's in the setting of $\mathbb{Z}[i]$, where we define $\Psi(x, c)$ to be the number of Gaussian integers $\alpha$ whose norm is at most $x$ and whose prime divisors have norm at most $x^c$.

While there is no simple function that gives the value of $\rho(u)$ for all $u$, there are numerous approximations, the accuracy of each dependent on the range of $u$ for which it is defined. For our purposes we will be interested in the following two approximations of $\rho$, both due to N.G. de Bruijn. The first approximation comes from [5, (1.6)],

**Theorem 2.1.3.** *For $u > 1$, let $\xi = \xi(u)$ be the positive root of the equation $e^{\xi} - 1 = u\xi$. Then,*

$$\rho(u) \sim \frac{e^{\gamma}}{\sqrt{2\pi u}} \exp\left(-\int_0^{\xi} \frac{se^s - e^s + 1}{s} ds\right)$$

*as $u \to \infty$, where $\gamma$ is Euler's constant.*

The next approximation is widely used and can be found in the same paper of de Bruijn [5, (1.8)]

**Theorem 2.1.4.** *For $u > 0$ we have the following,*

$$\rho(u) \;=\; \exp\left(-u\left(\ln(u)+\ln\ln(u)-1+\frac{\ln\ln(u)-1}{\ln(u)}+O\left(\frac{(\ln\ln(u))^2}{(\ln(u))^2}\right)\right)\right)$$

Our interest in $\rho$ may seem rather arbitrary given our setting is $\mathcal{O}_K$, where $K$ not necessarily $\mathbb{Q}$. However, as we shall see, many of the arguments made in $\mathcal{O}_K$ can be recast to the setting of the integers, in which case theorems 2.1.2. and 2.1.3. will be applicable. Furthermore, in the next section we define a function $\Psi_K(x, y)$ in $\mathcal{O}_K$, which is analogous to $\Psi(x, y)$ in $\mathbb{Z}$, and many approximations of this function also rely on $\rho$, see [15], [7], and [8].

## 2.2 Smooth Ideals in $\mathcal{O}_K$

Let $K$ be a finite extension of $\mathbb{Q}$ of degree $n$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. Analogous to our definition of smooth numbers, we define a $y$-smooth ideal of $\mathcal{O}_K$ as follows,

**Definition.** *For $y \geq 2$, we say that an ideal $I \subseteq \mathcal{O}_K$ is $y$-smooth if for every prime ideal $\mathfrak{p} \subseteq \mathcal{O}_K$ with $\mathfrak{p} \mid I$ we have $N(\mathfrak{p}) \leq y$. In other words, an ideal $I$ is said to be $y$-smooth if $I$ has no prime ideal factors with norm exceeding $y$.*

We define $\Psi_K(x, y)$ to be the number of ideals in $\mathcal{O}_K$ with norm at most $x$, having no prime ideal divisors with norm exceeding $y$. As with $\Psi(x, y)$, there is a great deal of literature regarding $\Psi_K(x, y)$, the interested reader should see [15], [1], and [8]. For our purposes, we will only concern ourselves with the following result due to P. Moree and C.L. Stewart [8, Thm. 2], and Canfield, Erdös, and Pomerance [2, Thm 3.1],

**Theorem 2.2.1.** *Let $K$ be a finite extension of $\mathbb{Q}$ of degree $n \geq 1$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. There exists a positive number $C_1 = C_1(K)$, which depends upon $K$, such that for all $x \geq 1$ and $u \geq 3$,*

$$\Psi_K(x, x^{1/u}) \ \geq \ x \ \exp\left(-u\left(\ln(u) + \ln\ln(u) - 1 + \frac{\ln\ln(u) - 1}{\ln(u)} + C_1\left(\frac{\ln\ln(u)}{\ln(u)}\right)^2\right)\right).$$

Thus far we have not given any motivation for our consideration of smooth numbers, or more generally, smooth ideals. The utility of such ideals, which will be discussed in the next section, lies in the fact that they give us a useful criterion for determining if an arbitrary sequence of ideals contains a square dependent subsequence.

## 2.3 Determining a Square Dependence

We begin by considering the following question: Suppose we have a sequence of ideals from $\mathcal{O}_K$, $I_1, \ldots, I_k$, and we know that this sequence has a square dependent subsequence. How would we go about determining this subsequence? Well, suppose we have two ideals $I, J \subseteq \mathcal{O}_K$. Since $\mathcal{O}_K$ is a Dedekind domain, we know that both $I$ and $J$ have unique factorizations (up to order) into the product of prime powers. Let $I = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_m^{e_m}$ and $J = \mathfrak{p}_1^{d_1} \cdots \mathfrak{p}_m^{d_m}$, where some of the $d_i$ and $e_i$ may be 0. Denote by $v(I)$ and $v(J)$, the vector consisting of the exponents in the prime ideal factorizations of $I$ and $J$, respectively. So, $v(I) = (e_1, e_2, \ldots, e_m)$ and $v(J) = (d_1, d_2, \ldots, d_m)$. Now, in order for $IJ$ to be a square, the exponent of each prime ideal in the prime ideal factorization of $IJ$ must be even, in other words, $v(IJ) \equiv (0, 0, 0, \ldots, 0) \bmod 2$. Since $v(IJ) = v(I) + v(J)$, this implies that the vectors $v(I)$ and $v(J)$ must be linearly dependent modulo 2. Therefore, if $I_1', \ldots, I_\ell'$ is a square dependent subsequence of $I_1, \ldots, I_k$, we have

$$\sum_{i=1}^{\ell} v(I_i^{'}) \equiv (0,0,0,\ldots,0) \bmod 2.$$

Thus, in order to determine the square dependent subsequence among the ideals $I_1,\ldots,I_k$, we need only determine the linear dependence among the vectors $v(I_1),\ldots,v(I_k)$. This is easily done utilizing methods from linear algebra. To see how, consider the following linear system of equations,

$$\begin{pmatrix} v(I_1)^T \bmod 2 & ,\cdots, & v(I_k)^T \bmod 2 \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

where $a_i \in \mathbb{Z}_2$. Solving for the $a_i$'s, we see that if $a_i = 1$ then $I_i$ is involved in the square dependency.

Not only does this simple argument from linear algebra give us a means of determining the square dependence among our sequence $x_1,\ldots,x_k$, it also gives us an means of determining if an arbitrary sequence of integers contains a square dependent subsequence. To see how, let $I_1, I_2, \ldots$ be an arbitrary sequence of ideals from $S(x)$. For a given $y \geq 2$, suppose we are able to find $\pi_K(y)+1$ terms of this sequence that are $y$-smooth, where $\pi_K(y)$ denotes the number of prime ideals in $\mathcal{O}_K$ with norm at most $y$. Since the exponent vector $v(I_i)$ will have dimension at most $\pi_K(y)$ for each $I_i$ that is $y$-smooth, and we have $\pi_K(y)+1$ such terms, we are guaranteed to have a square dependence by recalling the fact that $n+1$ vectors from an $n$ dimensional vector space must contain a linear dependence.

In the next section we prove the existence of a continuous function $\mu(c) > 0$, satisfying $\Psi(x,c) = \pi x \mu(c) + O(x/\ln(x))$ for all $c > 0$, where $\Psi(x,c)$ denotes the number of $x^c$-smooth Gaussian integers of norm at most $x$. This proof is analogous to the proof given by Ramaswami in [14], however, the result is not necessary to prove Theorem 1, and may be skipped.

## 2.4 The Dickman-de Bruijn Function for $\mathbb{Z}[i]$

We begin with some notation,

| | |
|---|---|
| $p, p_r$ : | any prime integer. |
| $\pi$ : | any Gaussian prime. |
| $S(x, p)$ : | the set of Gaussian integers whose norm is less than $x$,<br>divisible by $p$, and free of prime divisors greater than $p$. |
| $T(x, p)$ : | the set of Gaussian integers whose norm is less than $x$<br>and free of prime divisors greater than $p$. |
| $N(\alpha)$ : | the norm of the Gaussian integer $\alpha$ |
| $F(t)$ : | $\sum_{p \leq t} 1/p$ |

Next we define $\Psi(x, c)$ as follows,

$$\Psi(x, c) = \Big| \{\alpha \in \mathbb{Z}[i] : N(\alpha) < x,\, \pi \mid \alpha \Rightarrow N(\pi) < x^c\} \Big|.$$

That is, $\Psi(x, c)$ is the number of Gaussian integers with norm less than $x$, whose prime factors have norm at most $x^c$. We wish to prove the following theorem describing $\Psi(x, c)$,

**Theorem 2.4.1**. *A function $\mu(c)$ defined for all $c > 0$ exists such that*

1. *for any fixed c,*
$$\Psi(x, c) \;=\; \pi x \mu(c) + O\left(\frac{x}{\ln x}\right).$$
2. *$\mu(c) > 0$ and continuous for $c > 0$.*

Our previous notation would dictate that we consider the function $\Psi(x, x^{1/u})$, rather that $\Psi(x, c)$ with $c = 1/u$. However, aspects of the proof of Theorem 2.4.1 would make the notation $\Psi(x, x^{1/u})$ cumbersome, which is the reason for the change in notation. Before we prove Theorem 2.4.1, we will first prove a few preliminary lemmas.

**Lemma 1**. *For $c \geq 1$, Theorem 1 is true, and*

$$\Psi(x, c) \;=\; \pi x \mu(c) + O\,(\sqrt[3]{x}).$$

Proof. For $c \geq 1$, $\Psi(x, c)$ gives the number of Gaussian integers whose norm is less than $x$. This is equivalent to the number of lattice points in a circle of radius $\sqrt{x}$, which has been shown to be asymptotic to $\pi x + O(\sqrt[3]{x})$ by Sierpinski [16]. Since $O(\sqrt[3]{x}) = O(x/\ln x)$, Lemma 1 follows by taking $\mu(c) = 1$ for $c \geq 1$. ■

**Lemma 2**. *If $p_1 \neq p_2$, the sets $S(x, p_1)$ and $S(x, p_2)$ are distinct.*

Proof. Without loss of generality, we may assume $p_1 < p_2$. Let $\alpha \in S(x, p_1)$. Then the largest prime factor of $N(\alpha)$ is $p_1$, which implies $p_2$ does not divide $N(\alpha)$. Hence $\alpha \notin S(x, p_2)$. ■

**Lemma 3**.

$$\bigl|\, S(x, p) \,\bigr| \;=\; \begin{cases} 2\,\bigl|\, T(x/p, p) \,\bigr| - \bigl|\, T(x/p^2, p) \,\bigr| & \text{if } p \equiv 1 \ (4) \\ \\ \bigl|\, T(x/p^2, p) \,\bigr| & \text{if } p \equiv 3 \ (4) \end{cases}$$

Proof. Suppose $p \equiv 1$ (4), and let $\pi\overline{\pi} = p$. Define $A = \{\pi\beta \,:\, \beta \in T(x/p, p)\}$ and $B = \{\overline{\pi}\beta \,:\, \beta \in T(x/p, p)\}$. We want to show that $A \cup B = S(x, p)$. Let $\alpha \in A \cup B$. Then $\alpha = \pi\beta$ or $\alpha = \overline{\pi}\gamma$ for some $\beta, \gamma \in T(x/p,\, p)$. Without loss of generality, suppose $\alpha = \pi\beta$. Then $N(\alpha) = N(\pi)N(\beta) < p \cdot (x/p) = x$. Moreover, since $p \mid N(\alpha)$ and $p$ is the largest prime factor of $N(\alpha)$, we know $\alpha \in S(x, p)$. Hence, $A \cup B \subseteq S(x,\, p)$. Conversely, suppose $\alpha \in S(x, p)$. Then $p \mid N(\alpha)$, which implies either $\pi$ or $\overline{\pi}$ divides $\alpha$. Without loss of generality, suppose $\alpha = \pi\beta$. Then $N(\alpha) = N(\pi)N(\beta) < x$, which implies $N(\beta) < x/p$. Furthermore, since the largest prime factor of $N(\alpha)$ is $p$, we know that the largest prime factor of $N(\beta)$ is also $p$. This implies that $\beta \in T(x/p,\, p)$, which in turn implies $\alpha \in A \cup B$. Hence, $A \cup B = S(x,\, p)$. Therefore,

$$|\,S(x,p)\,| = |\,A \cup B\,| = |A| + |B| - |\,A \cap B\,| = 2\,|\,T(x/p,p)\,| - |\,A \cap B\,|.$$

Next we want to determine $|\,A \cap B\,|$. Let $C = \{\pi\overline{\pi}\beta \,:\, \beta \in T(x/p^2,p)\}$. Suppose $\alpha \in A \cap B$. Then it must be the case that $\pi\overline{\pi} \mid \alpha$, which implies $\alpha \in C$. Conversely, suppose $\alpha \in C$. Then we can write $\alpha = \pi\beta$ and $\alpha = \overline{\pi}\gamma$ for some $\beta, \gamma \in T(x/p,p)$. This implies that $\alpha \in A$ and $\alpha \in B$. Therefore $C = A \cap B$, and $|\,A \cap B\,| = |C| = |\,T(x/p^2,p)\,|$.

Suppose $p \equiv 3\ (4)$. Then $p$ is prime in $\mathbb{Z}[i]$, and the map $\sigma : T(x/p^2,p) \to S(x,p)$ defined by $\sigma(\alpha) = p\alpha$ is bijective. Hence $|\,S(x,p)\,| = |\,T(x/p^2,p)\,|$. ■

The next lemma will function as the induction hypothesis in our proof of Theorem 2.4.1.

**Lemma 4**. *If $c_1 \in (0,1]$, and Theorem 1 is true for $c \geq c_1$, then it is true for $c \geq c_1/(1+c_1)$.*

Proof. By hypothesis, $\mu(c)$ is defined, continuous, and positive for $c \geq c_1$. Furthermore, by considering Lemma 1 when $c \geq 1$ we have,

$$\Psi(r,c) \;=\; \begin{cases} \pi x\mu(c) + O\Big(\dfrac{x}{\ln x}\Big) & \text{for } c_1 \leq c < 1 \\ \\ \pi x\mu(c) + O\big(\sqrt[3]{x}\big) & \text{for } c \geq 1 \end{cases} \tag{2.1}$$

Let $c_2 = c_1/(1+c_1)$. Consider $d$ such that $c_2 \leq d \leq c_1$. We begin by showing that there is a function $\mu(d)$ such that $\Psi(x,d) = \pi x\mu(d) + O(x/\ln x)$. In order to do so, we consider the difference $\Psi(x,c_1) - \Psi(x,d)$. It follows from Lemma 2 that,

$$\Psi(x,c_1) - \Psi(x,d) \;=\; \sum_{\substack{x^d < p \leq x^{c_1} \\ p \equiv 1\,(4)}} |\,S(x,p)\,| \;+\; \sum_{\substack{x^{d/2} < p \leq x^{c_1/2} \\ p \equiv 3\,(4)}} |\,S(x,p)\,|. \tag{2.2}$$

Note that the second sum in (2.2) is over primes $p \equiv 3\ (4)$, hence $p$ is itself a Gaussian prime with $N(p) = p^2$. Therefore we must have $x^{d/2} < p \leq x^{c_1/2}$ in order for $x^d < N(p) \leq x^{c_1}$. By Lemma 3, (2.2) becomes

$$\sum_{\substack{x^d < p \leq x^{c_1} \\ p \equiv 1\,(4)}} \left( 2 \left| T\left(\frac{x}{p},\, p\right) \right| - \left| T\left(\frac{x}{p^2},\, p\right) \right| \right) + \sum_{\substack{x^{d/2} < p \leq x^{c_1/2} \\ p \equiv 3\,(4)}} \left| T\left(\frac{x}{p^2},\, p\right) \right| \tag{2.3}$$

Next, we note that the number of Gaussian integers whose norm in less than $x/p^2$ and free of prime divisors greater than $p$, is clearly less than or equal to the number of Gaussian integers whose norm is less than $x/p^2$, with the latter being equivalent to the number of lattice points in a circle of radius $\sqrt{x}/p$. Hence, we have the following,

$$\sum_{\substack{x^{d/2} < p \leq x^{c_1/2} \\ p \equiv 3\,(4)}} \left| T\left(\frac{x}{p^2},\, p\right) \right| \ll \sum_{\substack{x^{d/2} < p \leq x^{c_1/2} \\ p \equiv 3\,(4)}} \frac{x}{p^2} \leq \frac{x}{x^d} \cdot \left( \sum_{\substack{x^{d/2} < p \leq x^{c_1/2} \\ p \equiv 3\,(4)}} 1 \right)$$

$$\ll \frac{x}{x^d} \cdot \frac{x^{c_1/2}}{\ln x} = \frac{x^{1-d+(c_1/2)}}{\ln x}.$$

Recall that $c_1 \in (0, 1]$ and that $c_1/(c_1 + 1) \leq d \leq c_1$. Clearly $c_1/2 \leq c_1/(c_1 + 1) \leq d$, which implies $1 - d + c_1/2 \leq 1$. Hence,

$$\sum_{\substack{x^{d/2} < p \leq x^{c_1/2} \\ p \equiv 3\,(4)}} \left| T\left(\frac{x}{p^2},\, p\right) \right| \in O\left(\frac{x}{\ln x}\right) \tag{2.4}$$

Similarly,

$$\sum_{\substack{x^d < p \leq x^{c_1} \\ p \equiv 1\,(4)}} \left| T\left(\frac{x}{p^2},\, p\right) \right| \ll \sum_{\substack{x^d < p \leq x^{c_1} \\ p \equiv 1\,(4)}} \frac{x}{p^2} \leq \frac{x}{x^{2d}} \cdot \left( \sum_{\substack{x^d < p \leq x^{c_1} \\ p \equiv 1\,(4)}} 1 \right)$$

$$\ll \frac{x}{x^{2d}} \cdot \frac{x^{c_1}}{\ln x} = \frac{x^{c_1+1-2d}}{\ln x}.$$

Since $c_1 + 1 - 2d \leq (c_1^2 + 1)/(c_1 + 1) \leq 1$, we have

$$\sum_{\substack{x^d < p \leq x^{c_1} \\ p \equiv 1\,(4)}} \left| T\left(\frac{x}{p^2},\ p\right) \right| \ \in\ O\left(\frac{x}{\ln x}\right). \tag{2.5}$$

Taking (2.4) and (2.5) into account we can rewrite (2.3) as,

$$\sum_{\substack{x^d < p \leq x^{c_1} \\ p \equiv 1\,(4)}} 2 \left| T\left(\frac{x}{p},\ p\right) \right| \ +\ O\left(\frac{x}{\ln x}\right),$$

which, by our definition of $\Psi$, can written as,

$$\sum_{\substack{x^d < p \leq x^{c_1} \\ p \equiv 1(4)}} 2\,\Psi\left(\frac{x}{p},\ \frac{\ln p}{\ln x - \ln p}\right) \ +\ O\left(\frac{x}{\ln x}\right). \tag{2.6}$$

Furthermore, since $x^d < p$ we have $d(\ln x - \ln p) < \ln p\,(1-d)$, which implies $d/(1-d) < \ln p/(\ln x - \ln p)$. Hence,

$$\frac{\ln p}{\ln x - \ln p} > \frac{d}{1-d} > \frac{c_2}{1-c_2} = c_1.$$

Therefore, by hypothesis (2.1) we can write (2.6) as,

$$\sum_{\substack{x^d < p \leq x^{c_1} \\ p \equiv 1(4)}} \frac{2\pi x}{p}\,\mu\left(\frac{\ln p}{\ln x - \ln p}\right) + \sum_{\substack{x^d < p < \sqrt{x} \\ p \equiv 1(4)}} O\left(\frac{x/p}{\ln x/p}\right)$$
$$+ \sum_{\substack{\sqrt{x} \leq p \leq x^{c_1} \\ p \equiv 1(4)}} O\left(\sqrt[3]{x/p}\right) + O\left(\frac{x}{\ln x}\right). \tag{2.7}$$

It is a well known result in number theory that $F(t) = \sum_{p \leq t} 1/p = \ln\ln t + O(1)$. Therefore, considering just the second sum in (2.7) for a moment, and assuming $\sqrt{x}$ is not a prime congruent to 1 modulo 4, we get

$$\sum_{\substack{x^d < p < \sqrt{x} \\ p \equiv 1(4)}} O\left(\frac{x/p}{\ln x/p}\right) \ =\ O\left(\frac{x}{\ln x}\right) \sum_{\substack{x^d < p < \sqrt{x} \\ p \equiv 1(4)}} \frac{1}{p} \ =\ O\left(\frac{x}{\ln x}\right)(F(\sqrt{x}) - F(x^d))$$

$$= O\left(\frac{x}{\ln x}\right)\left(\ln\ln\sqrt{x} - \ln\ln x^d + O\left(1\right)\right)$$

$$= O\left(\frac{x}{\ln x}\right)\left(-\ln d - \ln 2 + O\left(1\right)\right) = O\left(\frac{x}{\ln x}\right).$$

If $\sqrt{x}$ is a prime congruent to 1 modulo 4, then the term $(F(\sqrt{x}) - F(x^d))$ above would be replaced by $(F(\sqrt{x}) - F(x^d) - 1/\sqrt{x})$. However, $-1/\sqrt{x} = O(1)$, leaving the result unchanged. Similarly, for the third sum in (2.7) we have

$$\sum_{\substack{\sqrt{x} \le p \le x^{c_1} \\ p \equiv 1(4)}} O\left(\sqrt[3]{x/p}\right) = O\left(\sqrt[3]{x}\right) \sum_{\substack{\sqrt{x} \le p \le x^{c_1} \\ p \equiv 1(4)}} \frac{1}{p^3} = O\left(\sqrt[3]{x}\right) = O\left(\frac{x}{\ln x}\right).$$

Hence, equation (2.7) can be reduced to

$$2\pi x \cdot \sum_{\substack{x^d < p \le x^{c_1} \\ p \equiv 1(4)}} \left(\frac{1}{p}\right) \mu\left(\frac{\ln p}{\ln x - \ln p}\right) + O\left(\frac{x}{\ln x}\right),$$

which can be written as,

$$2\pi x \cdot \sum_{\substack{x^d < p \le x^{c_1} \\ p \equiv 1(4)}} \mu\left(\frac{\ln p}{\ln x - \ln p}\right)\Big(F(p) - F(p-1)\Big) + O\left(\frac{x}{\ln x}\right). \tag{2.8}$$

We can further simplify (2.8) by noting,

$$\sum_{\substack{x^d < p \le x^{c_1} \\ p \equiv 1(4)}} \mu\left(\frac{\ln p}{\ln x - \ln p}\right)\Big(F(p) - F(p-1)\Big)$$
$$= \sum_{\substack{x^d < p \le x^{c_1} \\ p \equiv 1(4)}} \mu\left(\frac{\ln p}{\ln x - \ln p}\right)\Big(\ln\ln p - \ln\ln(p-1)\Big)$$
$$+ \sum_{\substack{x^d < p \le x^{c_1} \\ p \equiv 1(4)}} \mu\left(\frac{\ln p}{\ln x - \ln p}\right) O\left(1\right). \tag{2.9}$$

Since,

$$\sum_{\substack{x^d<p\leq x^{c_1}\\ p\equiv 1(4)}} \mu\left(\frac{\ln p}{\ln x-\ln p}\right) O\left(1\right) \;=\; O\left(\frac{x^{c_1}-x^d}{\ln x}\right) \;=\; O\left(\frac{x}{\ln x}\right),$$

we can write (2.9) as,

$$\sum_{\substack{x^d<p\leq x^{c_1}\\ p\equiv 1(4)}} \mu\left(\frac{\ln p}{\ln x-\ln p}\right)\Big(\ln\ln p-\ln\ln(p-1)\Big)+O\left(\frac{x}{\ln x}\right). \tag{2.10}$$

Hence (2.8) becomes,

$$2\pi x\cdot\sum_{\substack{x^d<p\leq x^{c_1}\\ p\equiv 1(4)}} \mu\left(\frac{\ln p}{\ln x-\ln p}\right)\Big(\ln\ln p-\ln\ln(p-1)\Big)+O\left(\frac{x}{\ln x}\right). \tag{2.11}$$

We would like to write the sum in (2.11) as a Riemann-Stieltjes integral, however, we must first convince ourselves that such an integral exists.

It was previously shown that $\ln p/(\ln x-\ln p)>c_1$ for all primes satisfying $x^d<p\leq x^{c_1}$. However, if $p=x^d$ we have $\ln p/(\ln x-\ln p)=d/(1-d)\geq c_2/(1-c_2)=c_1$. Hence $\ln p/(\ln x-\ln p)\geq c_1$ for all primes $p$ in the closed interval $[x^d,x^{c_1}]$, which implies $\mu(\ln p/(\ln x-\ln p))$ is continuous in the same interval. Furthermore, since $G(t)=\ln\ln t$ is a monotonically increasing function, the sum in (2.11) can be written as a Riemann-Stieltjes integral. Doing so gives us,

$$\begin{aligned}2\pi x\int_{x^d}^{x^{c_1}} \mu\left(\frac{\ln t}{\ln x-\ln t}\right) dG(t) \;+\; O\left(\frac{x}{\ln x}\right)&\\ =\; 2\pi x\int_{x^d}^{x^{c_1}} \mu\left(\frac{\ln t}{\ln x-\ln t}\right)\frac{dt}{t\ln t} \;+\; O\left(\frac{x}{\ln x}\right)&\end{aligned} \tag{2.12}$$

Performing the substitution $t = x^u$, (2.12) becomes

$$2\pi x \int_{d}^{c_1} \mu\left(\frac{u}{1-u}\right)\frac{du}{u} \;+\; O\left(\frac{x}{\ln x}\right). \tag{2.13}$$

Therefore,

$$\Psi(x,d) \;=\; \Psi(x,c_1) \;-\; 2\pi x \int_{d}^{c_1} \mu\left(\frac{u}{1-u}\right)\frac{du}{u} \;+\; O\left(\frac{x}{\ln x}\right) \tag{2.14}$$

which becomes,

$$\Psi(x,d) \;=\; \pi x\left(\mu(c_1) - 2\int_{d}^{c_1} \mu\left(\frac{u}{1-u}\right)\frac{du}{u}\right) \;+\; O\left(\frac{x}{\ln x}\right). \tag{2.15}$$

Since the integral in (2.15) and $\mu(c_1)$ are defined, the latter by hypothesis, we have shown that for any $c \geq c_1/(1+c_1)$, there exists a function $\mu$ such that $\Psi(x,c) = \pi x\mu(c) + O(x/\ln x)$. Next we will show that $\mu(c) > 0$ for all $c > c_1/(1+c_1)$.

Let $c_2 = c_1/(c_1+1) \leq d_1 < d_2 \leq c_1$. We want to show that $\mu(d_2) > 0$. By (2.15) we have,

$$\mu(d_2) - \mu(d_1) \;=\; \lim_{x\to\infty} \frac{\Psi(x,d_2) - \Psi(x,d_1)}{\pi x}$$

$$= \lim_{x\to\infty}\left[2\int_{d_1}^{d_2} \mu\left(\frac{u}{1-u}\right)\frac{du}{u} + O\left(\frac{1}{\ln x}\right)\right] \;=\; 2\int_{d_1}^{d_2} \mu\left(\frac{u}{1-u}\right)\frac{du}{u}.$$

Next we note that if $d_1 \leq u \leq d_2$, then

$$\frac{u}{1-u} \;\geq\; \frac{d_1}{1-d_1} \;\geq\; \frac{c_2}{1-c_2} \;=\; c_1.$$

Therefore, by hypothesis (2.1) we have $\mu(u/(1-u)) > 0$ for all $u \in [d_1, d_2]$, hence

$$\mu(d_2) - \mu(d_1) \;=\; 2\int\limits_{d_1}^{d_2} \mu\left(\frac{u}{1-u}\right)\frac{du}{u} \;>\; 0. \tag{2.16}$$

Note that (2.16) implies that $\mu$ is a strictly increasing function for $c \geq c_1/(c_1+1)$. Furthermore, we know that $\mu(c) \geq 0$ for $c \geq c_1/(c_1+1)$. To see this note that $\Psi(x,c) = ax$ for some positive $a \in \mathbb{R}$. If $\mu(c) = b < 0$, then by our definition we have $\Psi(x,c) = b\pi x + E$, where $E$ is the error term. This implies $E = (a - b\pi)x = O(x) \neq O(x/\ln x)$, a contradiction.

The fact that $\mu$ is a strictly increasing non-negative function for $c \geq c_1/(c_1+1)$ implies $\mu(d_2) \neq 0$. Hence, for any $c > c_1/(c_1+1)$ we have $\mu(c) > 0$. Note that we have not shown $\mu(c) > 0$ for $c = c_1/(c_1+1)$, we will return to this later. Next we want to show that $\mu(c)$ is continuous for all $c \in [c_1/(c_1+1),\, c_1]$.

Let $c \in [c_1/(c_1+1),\, c_1]$ and $\epsilon > 0$. Since $\ln(x)$ is continuous over the interval $[c_1/(c_1+1),\, c_1]$, we may choose a $\delta > 0$ such that for any $d \in [c_1/(c_1+1),\, c_1]$ satisfying $|c-d| < \delta$, we have $|\ln c - \ln d\,| < \epsilon/2$. Hence,

$$|\mu(c) - \mu(d)| \;=\; 2\left|\int\limits_{d}^{c} \mu\left(\frac{u}{1-u}\right)\frac{du}{u}\right| \;\leq\; 2\left|\int\limits_{d}^{c}\frac{du}{u}\right|$$

Where the last inequality follows from the fact that $\mu$ is a strictly increasing function for $c \geq c_1/(c_1+1)$, and by Lemma 1 we know $\mu(c) = 1$ for $c \geq 1 \geq c_1$. This gives us,

$$2\left|\int\limits_{d}^{c}\frac{du}{u}\right| \;=\; 2\,|\ln c - \ln d\,| \;<\; 2\cdot\frac{\epsilon}{2} \;=\; \epsilon.$$

Therefore, $\mu(c)$ is continuous for all $c \geq c_1/(c_1+1)$.

Now we return to the problem of showing that $\mu(c_1/(c_1+1)) > 0$. So far we have shown that if Theorem 2.4.1 holds for $c \geq c_1$, where $c_1 \in (0,1]$, then it holds for

$c > c_1/(c_1{+}1) = c_2$. Let $c^* \in (c_2, c_1\,)$, then Theorem 2.4.1 holds for $c \in (c^*/(c^*{+}1), c^*]$ by the current Lemma. Since $c_2 \in (c^*/(c^* + 1), c^*]$, we have $\mu(c_2) > 0$. ■

We are now in a position to prove Theorem 2.4.1.

Proof of Theorem 2.4.1. By Lemma 1, we know that Theorem 2.4.1 holds for $[1, \infty)$. Therefore, by Lemma 4 we know that Theorem 2.4.1 is true for $c \in [1/2, \infty)$. Suppose Theorem 2.4.1 holds for all $c \in [1/n, \infty)$ for all $n \leq k$. By Lemma 4, Theorem 2.4.1 also holds for all $c \in [k^*, \infty]$, where,

$$k^* \;=\; \frac{1/k}{1 + 1/k} \;=\; \frac{1}{k+1}.$$

The result follows by induction. ■

# Chapter 3

# The Square Threshold in $\mathcal{O}_K$

Let $K$ be a degree $n$ extension of $\mathbb{Q}$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. Recall that $L = \exp(\sqrt{\ln(x)\ln\ln(x)})$ and $S(x) = \{I \subseteq \mathcal{O}_K : N(I) \leq x\}$. The goal of this chapter is to prove the following theorem,

**Theorem 1.** *Let $K$ be a degree $n$ extension of $\mathbb{Q}$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. Let $\Omega$ be the smallest integer such that the first $\Omega$ terms of an ideal sequence generated by randomly and independently selecting ideals with uniform probability from the set $S(x)$ contains a square dependent subsequence. Then for any $\epsilon > 0$ the probability that $\Omega$ is in the interval*

$$\left(L^{\sqrt{2}-\epsilon},\, L^{\sqrt{2}+\epsilon}\right)$$

*tends to 1 as $x \to \infty$.*

We will prove Theorem 1 by breaking it into two parts. For $\epsilon > 0$, we will first prove the upper bound by showing the probability that $\Omega < (1+\epsilon)L^{\sqrt{2}}$ tends to 1 as $x \to \infty$. We will then prove the lower bound by showing the probability that $\Omega \leq L^{\sqrt{2}-\epsilon}$ tends to 0 as $x \to \infty$. Throughout both proofs $\mathfrak{p}, \mathfrak{p}_i \subseteq \mathcal{O}_K$ will denote prime ideals and $p, p_i \in \mathbb{Z}$ will denote prime integers, for all $i \in \mathbb{N}$.

Before we begin, we need to take note of an important theorem due to Edmund Landau [8], which serves as the analogue of Prime Number Theorem. It gives us an asymptotic expression for the number of prime ideals with norm below a specified bound.

**Theorem 3.1.1** (Prime Ideal Theorem) *Let $\pi_K(x)$ denote the number of prime ideals in $\mathcal{O}_K$ with norm at most $x$. Then,*

$$\pi_K(x) \sim \frac{x}{\ln(x)}$$

Furthermore, it follows from Theorem 3.1.1 that for $x \geq 2$,

$$\pi_K(x) = \mathrm{Li}(x) + O\left(xe^{-c\sqrt{\ln(x)}}\right)$$

where $c > 0$ is a constant depending on $K$ only, and $\mathrm{Li}(x) = \int_2^x 1/\ln(t)\,dt$ [8, (12)]. We are now ready to begin our proof of the upper bound given in Theorem 1.

## 3.1 The Upper Bound

We begin our proof with a generalization of a result due to N.G. de Bruijn [10, Thm. 2].

**Theorem 3.1.2.** *Let $K$ be a finite extension of $\mathbb{Q}$ of degree $n \geq 2$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. If $c > 1$ is constant, we have for $x \geq 2$ and $(\ln(x))^c \leq y \leq x$,*

$$\ln(\Psi_K(x,y)) \;\leq\; \ln(x\rho(u)) + \frac{\ln(1+u)}{2} + O(\ln\ln(y)) + O\left(\frac{(\ln(x))^2}{y}\right) + O(R)$$

*where*

$$R = \int_1^{\ln(y)} e^{s(\alpha/\ln(y))} V(e^s)\,ds; \quad u = \frac{\ln(x)}{\ln(y)}; \;\; \alpha = \ln(u) + \ln\ln(u+1)$$

*and $V$ is a function connected with the error term in the Prime Ideal Theorem.*

Proof. Let $Y = \{I \subseteq \mathcal{O}_K : \mathfrak{p} \mid I \Rightarrow N(\mathfrak{p}) \leq y\}$. That is, let $Y$ be the set of ideals in $\mathcal{O}_K$ that are $y$-smooth. Then for any $\eta > 0$ we have,

$$\begin{aligned}\Psi_K(x,y) \;&=\; \sum_{\substack{N(I)\leq x\\ I\in Y}} 1 \;\leq\; \sum_{\substack{N(I)\leq x\\ I\in Y}} \left(\frac{x}{N(I)}\right)^{\eta} \leq\; \sum_{I\in Y}\left(\frac{x}{N(I)}\right)^{\eta}\\ &=\; x^{\eta}\sum_{I\in Y}\frac{1}{N(I)^{\eta}} \;=\; x^{\eta}\prod_{N(\mathfrak{p})\leq y}\left(1-N(\mathfrak{p})^{-\eta}\right)^{-1}.\end{aligned}$$

That is, for all $\eta > 0$ we have,

$$\Psi_K(x,y) \;\leq\; x^{\eta}\prod_{N(\mathfrak{p})\leq y}\left(1-N(\mathfrak{p})^{-\eta}\right)^{-1},$$

which implies,

$$\ln(\Psi_K(x,y)) \;\leq\; \eta\ln(x) + \sum_{N(\mathfrak{p})\leq y}\ln\left(\left(1-N(\mathfrak{p})^{-\eta}\right)^{-1}\right). \tag{3.1}$$

We will estimate the sum in (3.1) by the following integral,

$$\int_e^y \ln\left((1-t^{-\eta})^{-1}\right) d\mathrm{Li}(t). \tag{3.2}$$

In order to do so, we must determine the amount of error that will be incurred by such an estimate. By the Prime Ideal Theorem, for $y \geq 2$ we have

$$\pi_K(y) = \mathrm{Li}(y) + O\left(ye^{-c\sqrt{\ln(y)}}\right). \tag{3.3}$$

where, for convenience, we may take $\mathrm{Li}(y) = \int_e^y 1/\ln(t)\,dt$. Noting that,

$$\int_e^y e^{-c\sqrt{\ln(t)}} - \frac{ce^{-c\sqrt{\ln(t)}}}{2\sqrt{\ln(y)}}dt \;=\; ye^{-c\sqrt{\ln(y)}} - e^{-c+1} \;=\; O\left(ye^{-c\sqrt{\ln(y)}}\right),$$

we can write (3.3) as,

$$\pi_K(y) = \mathrm{Li}(y) + O\left(\int_e^y e^{-c\sqrt{\ln(t)}} - \frac{ce^{-c\sqrt{\ln(t)}}}{2\sqrt{\ln(y)}} dt\right). \tag{3.4}$$

For ease of notation, we will denote $e^{-c\sqrt{\ln(t)}} - ce^{-c\sqrt{\ln(t)}}/2\sqrt{\ln(y)}$ by $V(t)$. Let $E$ denote the error incurred by approximating of the sum in (3.1) by (3.2). Then,

$$E = \sum_{N(\mathfrak{p})\leq y} \ln\left(\left(1 - N(\mathfrak{p})^{-\eta}\right)^{-1}\right) - \int_e^y \ln\left((1-t^{-\eta})^{-1}\right) d\mathrm{Li}(t).$$

Note that we can write the above sum as,

$$\ln\left((1-2^{-\eta})^{-1}\right) + \sum_{e\leq n\leq y} \ln\left((1-n^{-\eta})^{-1}\right)(\pi_K(n) - \pi_K(n-1)),$$

where $n \in \mathbb{N}$. Since the function $\ln\left((1-n^{-\eta})^{-1}\right)$ is continuous over $[e, y]$ with respect to $n$, and $\pi_K$ is monotonically increasing, we can write the above sum in terms of a Riemann-Stieltjes integral. Doing so we get,

$$\sum_{e\leq n\leq y} \ln\left((1-n^{-\eta})^{-1}\right)(\pi_K(n) - \pi_K(n-1)) = \int_e^y \ln\left((1-t^{-\eta})^{-1}\right) d\pi_K(t)$$

Hence,

$$\begin{aligned} E &= \ln\left((1-2^{-\eta})^{-1}\right) + \int_e^y \ln\left((1-t^{-\eta})^{-1}\right) d(\pi_K(t) - \mathrm{Li}(t)) \\ &= O(1) + \int_e^y \ln\left((1-t^{-\eta})^{-1}\right) d(\pi_K(t) - \mathrm{Li}(t)) \end{aligned} \tag{3.5}$$

For ease of notation, let $f(t) = \ln\left((1-t^{-\eta})^{-1}\right)$. Performing integration by parts on the integral in (3.5), we get

$$\int_e^y f(t) d\left(\pi_K(t) - \mathrm{Li}(t)\right) = f(t)\left(\pi_K(t) - \mathrm{Li}(t)\right)\Big|_e^y - \int_e^y f'(t)\left(\pi_K(t) - \mathrm{Li}(t)\right) dt$$

$$= \; f(y)\left(\pi_K(y) - \mathrm{Li}(y)\right) - f(e)\left(\pi_K(e) - \mathrm{Li}(e)\right) - \int_e^y f'(t)\, O\left(\int_e^t V(s)\, ds\right) dt \tag{3.6}$$

with the last equality following from (3.4). Considering just the integral in (3.6) for a moment, we note that

$$= \; \int_e^y f'(t)\, O\left(\int_e^t V(s)\, ds\right) dt \; = \; O\left(\int_e^y -f'(t)\left(\int_e^t V(s)\, ds\right) dt\right).$$

Performing integration by parts on the right hand side we get,

$$O\left(\int_e^y -f'(t)\left(\int_e^t V(s)\, ds\right) dt\right) \; = \; O\left(-f(t)\left(\int_e^t V(s)\, ds\right)\bigg|_e^y + \int_e^y f(t)V(t)dt\right)$$

$$= \; O\left(-f(y)\int_e^y V(s)\, ds + \int_e^y f(t)V(t)dt\right) \; = \; O\left(\int_e^y \left(f(t) - f(y)\right) V(t)dt\right)$$

Since $f(t)$ is a positive, monotonically decreasing function over the interval $[e, y]$, we know that

$$\int_e^y \left(f(t) - f(y)\right) V(t)dt \; = \; O\left(\int_e^y f(t)V(t)dt\right).$$

Thus,

$$\int_e^y f'(t)\, O\left(\int_e^t V(s)\, ds\right) dt \; = \; O\left(\int_e^y f(t)V(t)dt\right). \tag{3.7}$$

Next we note that,

$$f(y)\left(\pi_K(y) - \mathrm{Li}(y)\right) - f(e)\left(\pi_K(e) - \mathrm{Li}(e)\right) \; = \; O\left(\int_e^y f(y)V(t)dt\right) + O(1),$$

and since, $f(y) \leq f(t)$ for all $t \in [e, y]$, we have

$$O\left(\int_e^y f(y)V(t)dt\right) \; = \; O\left(\int_e^y f(t)V(t)dt\right).$$

Therefore, by (3.7) and the above equality, the right hand side of (3.6) becomes,

$$O\left(\int_e^y f(t)V(t)dt\right) + O(1),$$

which implies that,

$$E \;=\; O\left(\int_e^y f(t)V(t)dt\right) + O(1) \;=\; O\left(\int_e^y \ln\left((1-t^{-\eta})^{-1}\right)V(t)dt\right) + O(1).$$

Therefore, if we replace the sum in (3.1) with the integral in (3.2), and account for the resulting error we have the following,

$$\begin{aligned}\ln(\Psi_K(x,y)) \;\le\; \eta\ln(x) + \int_e^y \ln\left(\left(1-t^{-\eta}\right)^{-1}\right)\left(\frac{1}{\ln(t)}\right)dt \\ +O\left(\int_e^y \ln\left((1-t^{-\eta})^{-1}\right)V(t)dt\right) + O(1). \quad (3.8)\end{aligned}$$

Let $c \in \mathbb{R}$ such that $c > 1$ and restrict the values of $x$ and $y$ by $x > e^c$ and

$$\exp\left(\frac{c\ln\ln(x) - c\ln(c)}{1 - c/\ln(x)}\right) \;\le\; y < x \qquad (3.9)$$

Recall the function $\xi = \xi(u)$ from Theorem 3.1.1., which is defined to be the positive root of the equation $e^\xi - 1 = \xi u$ with $u = \ln(x)/\ln(y) > 1$. Let $\tau = 1 - \xi/\ln(y)$. For reasons that will be made clear later on, we would like to let $\eta = \tau$ in (3.8), however, in order to do so we must first show $\tau > 0$ for $x$ and $y$ satisfying $x > e^c$ and (3.9). Let us fix $x$. Note that as $y$ increases through the interval given by (3.9), $u = (e^\xi - 1)/\xi$ decreases, hence, $\xi$ must also be decreasing. Next we want to determine how large $\xi$ can be for $y$ in (3.9). Plugging in the minimum value of $y$ over (3.9) we get $u = (\ln(x) - c)/(c\ln\ln(x) - c\ln(c))$, which implies $\xi = \ln\ln(x) - \ln(c)$. Hence, $0 < \xi \le \ln\ln(x) - \ln(c)$. Since $\tau$ is minimized when $\xi$ is maximized, we have

$$\begin{aligned}\tau \;=\; 1 - \xi/\ln(y) \;=\; 1 - (e^\xi - 1)/\ln(x) \\ \ge\; 1 - (e^{\ln\ln(x) - \ln(c)} - 1)/\ln(x) \;=\; 1 - \frac{1}{c} + \frac{1}{\ln(x)}.\end{aligned}$$

Since $c > 1$, we have $\tau > 0$. Hence, we may let $\eta = \tau$ in (3.8), which gives us,

$$\ln(\Psi_K(x,y)) \;\leq\; \ln(x) - \xi u + \int_e^y \ln\left(\left(1 - t^{-\tau}\right)^{-1}\right)\left(\frac{1}{\ln(t)} + O\bigl(V(t)\bigr)\right) dt + O(1). \tag{3.10}$$

Next we would like to show $\ln\left((1-t^{-\tau})^{-1}\right) = t^{-\tau} + O(t^{-2\tau})$. To do so, we note the Mercator series expansion for $\ln(1+x)$, which is valid for $-1 < x \leq 1$,

$$\ln(1+x) \;=\; \sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}.$$

With respect the integral in (3.10), we see that $t > 1$, and since we have previously shown that $\tau > 0$, we know that $t^\tau > 1$. Therefore, $-1 < -1/t^\tau < 0$. Letting $x = -1/t^\tau$ in the above Mercator series expansion for $\ln(1+x)$ we get,

$$\ln\left(\frac{t^\tau - 1}{t^\tau}\right) \;=\; -\sum_{n=1}^{\infty}\frac{1}{nt^{n\tau}}.$$

Finally we note that,

$$\ln\left(\frac{1}{1 - 1/t^\tau}\right) \;=\; \ln\left(\frac{t^\tau}{t^\tau - 1}\right) \;=\; -\ln\left(\frac{t^\tau - 1}{t^\tau}\right) \;=\; \sum_{n=1}^{\infty}\frac{1}{nt^{n\tau}},$$

which gives us $\ln\left((1-t^{-\tau})^{-1}\right) = t^{-\tau} + O(t^{-2\tau})$. Therefore, (3.10) becomes

$$\begin{aligned}\ln(\Psi_K(x,y)) \;\leq\;& \ln(x) - \xi u + \int_e^y \frac{1}{t^\tau \ln(t)}dt + O\left(\int_e^y \frac{V(t)}{t^\tau}dt\right) \\ &+ O\left(\int_e^y \frac{1}{t^{2\tau}\ln(t)}dt\right) + O\left(\int_e^y \frac{V(t)}{t^{2\tau}}dt\right) + O(1). \end{aligned} \tag{3.11}$$

Since $O\left(\int_e^y V(t)\,dt\right)$ is defined to be the error term in the Prime Ideal Theorem, that is, $\pi(y) = \int_e^y 1/\ln(t) + O(\int_e^y V(t))$, we have $V(t) = O(1/\ln(t))$. Hence, we can reduce

(3.11) further, giving us

$$\begin{aligned}\ln(\Psi_K(x,y)) \;\leq\; \ln(x) - \xi u + \int_e^y \frac{1}{t^\tau \ln(t)} dt + O\left(\int_e^y \frac{V(t)}{t^\tau} dt\right) \\ + O\left(\int_e^y \frac{1}{t^{2\tau}\ln(t)} dt\right) + O(1). \quad (3.12)\end{aligned}$$

Considering just the first integral of (3.12) for a moment, we make the substitution $t = y^{s/\xi}$. Doing so we get $s = \xi \ln(t)/\ln(y)$ and $ds = \xi/(y^{s/\xi}\ln(y))$, which gives us

$$\begin{aligned}\int_e^y \frac{1}{t^\tau \ln(t)} dt \;=\; \int_{1-\tau}^{\xi} \left(\frac{1}{y^{s/\xi}}\right)^{\tau} \frac{\xi}{s\ln(y)} \left(\frac{y^{s/\xi}\ln(y)}{\xi}\right) ds \;=\; \int_{1-\tau}^{\xi} \left(\frac{1}{y^{s/\xi}}\right)^{\tau-1} \frac{1}{s}\, ds \\ \int_{1-\tau}^{\xi} \left(\frac{1}{y^{s/\xi}}\right)^{-\xi/\ln(y)} \frac{1}{s}\, ds \;=\; \int_{1-\tau}^{\xi} y^{s/\ln(y)} \frac{1}{s}\, ds \;=\; \int_{1-\tau}^{\xi} \frac{e^s}{s}\, ds. \quad (3.13)\end{aligned}$$

Recall from Theorem 2.1.3, that

$$\rho(u) \sim \frac{e^\gamma}{\sqrt{2\pi u}} \exp\left(-\int_0^\xi \frac{se^s - e^s + 1}{s} ds\right),$$

as $u \to \infty$, where $\xi$ be the positive root of the equation $e^\xi - 1 = u\xi$ and $\gamma$ is Euler's constant. Hence,

$$\ln(\rho(u)) \;=\; O(1) - \frac{1}{2}\ln(u) - \int_0^{1-\tau} \frac{se^s - e^s + 1}{s} ds - \int_{1-\tau}^{\xi} \frac{se^s + 1}{s} ds + \int_{1-\tau}^{\xi} \frac{e^s}{s} ds$$

which implies,

$$\begin{aligned}\int_{1-\tau}^{\xi} \frac{e^s}{s} ds \;&=\; \ln(\rho(u)) + \frac{\ln(u)}{2} + O(1) + \int_0^{1-\tau} \frac{se^s - e^s + 1}{s} ds + \int_{1-\tau}^{\xi} \frac{se^s + 1}{s} ds \\ &=\; \ln(\rho(u)) + \frac{\ln(u)}{2} + O(1) + \int_0^{\xi} e^s\, ds + \int_0^{1-\tau} \frac{1 - e^s}{s}\, ds + \int_{1-\tau}^{\xi} \frac{1}{s}\, ds \\ &=\; \ln(\rho(u)) + \frac{\ln(u)}{2} + O(1) + e^\xi - 1 + \int_0^{1-\tau} \frac{1 - e^s}{s}\, ds + \ln(\xi) - \ln(1-\tau) \\ &=\; \ln(\rho(u)) + \frac{\ln(u)}{2} + O(1) + u\xi + \int_0^{1-\tau} \frac{1 - e^s}{s}\, ds + \ln(\xi) - \ln(\xi/\ln(y)) \\ &=\; \ln(\rho(u)) + \frac{\ln(u)}{2} + O(1) + u\xi + \int_0^{1-\tau} \frac{1 - e^s}{s}\, ds + \ln\ln(y)\end{aligned}$$

Next we note that $(1/2)\ln(u+1) = (1/2)\ln(u) + O(1)$, giving us

$$\int_{1-\tau}^{\xi} \frac{e^s}{s} ds \;=\; \ln(\rho(u)) + \frac{\ln(u+1)}{2} + O(1) + u\xi + \int_0^{1-\tau} \frac{1-e^s}{s}\, ds + \ln\ln(y).$$

By (3.9) and the fact that $0 < \xi < \ln\ln(x) - \ln(c)$, we have $1 - \tau = \xi/\ln(y) \leq 1/c - 1/\ln(x) < 1$ for all $x$. Furthermore, by noting that $\lim_{s\to 0}(1-e^s)/s = 1$, we have,

$$\left|\int_0^{1-\tau} \frac{1-e^s}{s}\, ds\right| \;<\; \left|\int_0^{1} \frac{1-e^s}{s}\, ds\right| \;=\; O(1)$$

Hence,

$$\int_{1-\tau}^{\xi} \frac{e^s}{s} ds \;=\; \ln(\rho(u)) + \frac{\ln(u+1)}{2} + O(1) + u\xi + \ln\ln(y).$$

Plugging the above into (3.12) we get,

$$\begin{aligned}\ln(\Psi_K(x,y)) \;\leq\; & \ln(x) + \ln(\rho(u)) + \ln\ln(y) + \frac{\ln(u+1)}{2} \\ & + \; O\left(\int_e^y \frac{1}{t^{2\tau}\ln(t)} dt\right) + O\left(\int_e^y \frac{V(t)}{t^{\tau}} dt\right) + O(1). \quad (3.14)\end{aligned}$$

Next we will consider the error terms of (3.14). We begin by letting $t = e^s$ in the first error term,

$$O\left(\int_e^y \frac{1}{t^{2\tau}\ln(t)}\, dt\right) \;=\; O\left(\int_1^{\ln(y)} \frac{e^s}{e^{2\tau s}s}\, ds\right) \;=\; O\left(\int_1^{\ln(y)} \frac{e^{(1-2\tau)s}}{s}\, ds\right)$$

Recall that $0 < \xi \leq \ln\ln(x) - \ln(c)$ for $y$ in the interval (3.9), where $c > 1$ is constant. This implies that $\ln\ln(x) \geq \xi$. Letting $\omega = 2(\ln\ln(x)/\ln(y)) - 1$, we see that $\omega \geq 2(\xi/\ln(y)) - 1 = 1 - 2\tau$. Hence,

$$\int_1^{\ln(y)} \frac{e^{(1-2\tau)s}}{s}\, ds \;=\; O\left(\int_1^{\ln(y)} \frac{e^{\omega s}}{s}\, ds\right)$$

Note that if $\omega < 1/\ln(y)$, then $e^{\omega s} < e$ for all $s \in [1, \ln(y)]$. In other words, if $\omega < 1/\ln(y)$ then $e^{\omega s} = O(1)$, which gives us

$$\int_1^{\ln(y)} \frac{e^{\omega s}}{s}\, ds \;=\; O\left(\int_1^{\ln(y)} \frac{1}{s}\, ds\right) \;=\; O\Big(\ln\ln(y)\Big).$$

Suppose $\omega > 1/\ln(y)$. If $\omega \leq 1$ we can break up the integral as follows,

$$\int_1^{\ln(y)} \frac{e^{\omega s}}{s}\, ds \;=\; \int_1^{1/\omega} \frac{e^{\omega s}}{s}\, ds + \int_{1/\omega}^{\ln(y)} \frac{e^{\omega s}}{s}\, ds.$$

The first integral is $O(\ln\ln(y))$ by the previous argument. Letting $\omega s = z$ in the second integral we get

$$\int_{1/\omega}^{\ln(y)} \frac{e^{\omega s}}{s}\, ds \;=\; \int_1^{\omega\ln(y)} \frac{e^{z}}{z}\, dz \;=\; O\left(e^{\omega\ln(y)}\right),$$

Next, suppose $\omega > 1$. Since $1/\omega > 0$ and $e^{\omega s}/s > 0$ for all $s > 0$, we have

$$\int_1^{\ln(y)} \frac{e^{\omega s}}{s}\, ds \;\leq\; \int_{1/\omega}^{\ln(y)} \frac{e^{\omega s}}{s}\, ds \;=\; O\left(e^{\omega\ln(y)}\right).$$

Hence, the first error error term of (3.14) can be written as,

$$\int_e^{y} \frac{1}{t^{2\tau}\ln(t)}\, dt \;=\; O\Big(\ln\ln(y)\Big) + O\left(e^{\omega\ln(y)}\right)$$

$$= \; O\Big(\ln\ln(y)\Big) + O\left(\frac{(\ln(x))^2}{y^2}\right) \;=\; O\Big(\ln\ln(y)\Big) + O\left(\frac{(\ln(x))^2}{\ln(y)}\right). \qquad (3.15)$$

with the last equality following from that fact that $y > 1$. To get a handle on the second error term of (3.14), we first let $t = e^s$, giving us

$$\int_e^{y} \frac{V(e^s)}{e^{s\tau}}\, dt \;=\; \int_1^{\ln(y)} \frac{V(e^s)}{e^{s(1-\xi/\ln(y))}}\, ds \;=\; O\left(\int_1^{\ln(y)} V(e^s) e^{s(\xi/\ln(y))}\, ds\right). \qquad (3.16)$$

Next we want to replace $\xi$ in (3.16) with $\ln(u) + \ln\ln(u+1)$. Note that since $u = (e^{\xi} - 1)/\xi$ and $u + 1 = (e^{\xi} + \xi - 1)/\xi$, we have

$$\ln(u) + \ln\ln(u+1) \;=\; \ln(e^{\xi}-1) - \ln(\xi) + \ln\ln\left(\frac{e^{\xi}+\xi-1}{\xi}\right).$$

However, since

$$\lim_{\xi\to\infty} \frac{\ln(e^{\xi}-1)}{\xi} \;=\; 1 \quad \text{and} \quad \lim_{\xi\to\infty}\left(\ln\ln\left(\frac{e^{\xi}+\xi-1}{\xi}\right) - \ln(\xi)\right) \;=\; 0$$

we have $\ln(e^{\xi}-1) \sim \xi$ and $\ln\ln((e^{\xi}+\xi-1)/\xi) - \ln(\xi) = o(1)$. Hence,

$$\xi = \ln(u) + \ln\ln(u+1) + O(1).$$

Let $\alpha = \ln(u) + \ln\ln(u+1)$. Then,

$$\begin{aligned}\int_1^{\ln(y)} V(e^s)e^{s(\xi/\ln(y))}\,ds \;&=\; \int_1^{\ln(y)} V(e^s)e^{s((\alpha+O(1))/\ln(y))}\,ds \\ &=\; O\left(\int_1^{\ln(y)} V(e^s)e^{s(\alpha/\ln(y))}\,ds\right). \qquad (3.17)\end{aligned}$$

Therefore, by plugging in (3.15) and (3.17) for the first and second error terms of (3.14), respectively, we get

$$\begin{aligned}\ln(\Psi_K(x,y)) \;\le\; & \ln(x) + \ln(\rho(u)) + \frac{\ln(u+1)}{2} \\ & + O(\ln\ln(y)) + O\left(\frac{(\ln(x))^2}{\ln(y)}\right) + O\left(\int_1^{\ln(y)} V(e^s)e^{s(\alpha/\ln(y))}\,ds\right). \qquad (3.18)\end{aligned}$$

Note that (3.18) is the desired expression in Theorem 3.1.2. All that remains to be shown is that (3.18) holds in the range specified in Theorem 3.1.2. Note that for $x \ge 2$ satisfying $\ln(x)/\ln\ln(x) \ge c/\ln(c)$ for the constant $c > 1$, we have

$$\exp\left(\frac{c\ln\ln(x) - c\ln(c)}{1 - c/\ln(x)}\right) \;\le\; \ln(x)^c. \qquad (3.19)$$

Let $c_1 \in \mathbb{R}$ be a constant for which (3.19) is satisfied for all $x > c_1$. Then we have shown for $x > c_1$ with $(\ln(x))^c \le y \le x$ the result holds.

Finally, we note that for $2 \leq x \leq c_1$ with $(\ln(x))^c \leq y \leq x$, each of the terms on the right hand side of (3.18) are finite, and $\ln(x)^2/y \geq \ln(2)^2/2 > 0$, we see that the result follows trivially for $2 \leq x \leq c_1$ with $(\ln(x))^c \leq y \leq x$ since the $O(\ln(x)^2/y)$ term can encompass any possible error.

■

Now that we have an upper bound for $\Psi_K(x,y)$, or more accurately an upper bound for $\ln(\Psi_K(x,y))$, we will use Theorems 3.1.2 and 2.2.1 to sandwich $\Psi_K(x,y)$ and allow us determine an asymptotic equality. This result is a generalization of that given by Pomerance [13, Thm. 2.1] in the case when $K = \mathbb{Q}$.

**Corollary 3.1.3.** *Let $y = x^{1/u}$. If $\epsilon > 0$ is arbitrary, for $x \geq 2$ and $3 \leq u \leq (1-\epsilon)\ln(x)/\ln\ln(x)$, we have*

$$\Psi_K(x,y) \;=\; \frac{|S(x)|}{u^{u(1+o(1))}}$$

*as $u \to \infty$.*

Proof. First we will show that $\Psi_K(x,y) = x/u^{u(1+o(1))}$ as $u \to \infty$. By Theorem 2.2.1, we have

$$\Psi_K(x,x^{1/u}) \;\geq\; x\,\exp\left(-u\left(\ln(u)+\ln\ln(u)-1+\frac{\ln\ln(u)-1}{\ln(u)}+C_1\left(\frac{\ln\ln(u)}{\ln(u)}\right)^2\right)\right)$$

$$=\; x\,\exp\left(-u\ln(u)\left(1+\frac{\ln\ln(u)}{\ln(u)}-\frac{1}{\ln(u)}+\frac{\ln\ln(u)-1}{(\ln(u))^2}+C_1\frac{(\ln\ln(u))^2}{(\ln(u))^3}\right)\right)$$

$$=\; x\,u^{\left(-u\left(1+\frac{\ln\ln(u)}{\ln(u)}-\frac{1}{\ln(u)}+\frac{\ln\ln(u)-1}{(\ln(u))^2}+C_1\frac{(\ln\ln(u))^2}{(\ln(u))^3}\right)\right)} \;=\; \frac{x}{u^{u(1+o(1))}}$$

as $u \to \infty$.

By Theorems 3.1.2 and 2.1.4, we have

$$\begin{aligned}
&\ln(\Psi_K(x, x^{1/u})) \\
&\quad \leq \ \ln\left(x \exp\left(-u\left(\ln(u) + \ln\ln(u) - 1 + \frac{\ln\ln(u) - 1}{\ln(u)} + C_1\left(\frac{\ln\ln(u)}{\ln(u)}\right)^2\right)\right)\right) \\
&\qquad\qquad\qquad + \frac{\ln(u+1)}{2} + O(\ln\ln(y)) + O\left(\frac{(\ln(x))^2}{y}\right) + O(R).
\end{aligned}$$

Exponentiating both sides of the inequality, we may apply the same argument as above giving us

$$\Psi_K(x, x^{1/u}) \ \leq \ \frac{x}{u^{u(1+o(1))}} \exp\left(\frac{\ln(u+1)}{2} + O(\ln\ln(y)) + O\left(\frac{(\ln(x))^2}{y}\right) + O(R)\right).$$

as $u \to \infty$. Hence, it will suffice to show that

$$\exp\left(\frac{\ln(u+1)}{2} + O(\ln\ln(y)) + O\left(\frac{(\ln(x))^2}{y}\right) + O(R)\right) \ = \ u^{-u(o(1))}$$

as $u \to \infty$. We have,

$$\begin{aligned}
&\exp\left(\frac{\ln(u+1)}{2} + O(\ln\ln(y)) + O\left(\frac{(\ln(x))^2}{y}\right) + O(R)\right) \\
= \ &\exp\left(-u\ln(u)\left(-\frac{\ln(u+1)}{u\ln(u)} + O\left(\frac{\ln\ln(y)}{u\ln(u)}\right) + O\left(\frac{(\ln(x))^2}{yu\ln(u)}\right) + O\left(\frac{R}{u\ln(u)}\right)\right)\right) \\
= \ &u^{-u\left(-\frac{\ln(u+1)}{u\ln(u)} + O\left(\frac{\ln\ln(y)}{u\ln(u)}\right) + O\left(\frac{(\ln(x))^2}{yu\ln(u)}\right) + O\left(\frac{R}{u\ln(u)}\right)\right)}.
\end{aligned}$$

Clearly, $-\ln(u+1)/u\ln(u) = o(1)$ as $u \to \infty$, so all that remains to be shown is that the subsequent terms are also $o(1)$. Note that since $y = x^{1/u}$, we have $\ln\ln(y) = \ln\ln(x) - \ln(u)$ and $u = \ln(x)/\ln(y)$. Therefore,

$$\frac{\ln\ln(y)}{u\ln(u)} \ = \ \frac{\ln\ln(x) - \ln(u)}{u\ln(u)} \ = \ \frac{\ln\ln(x)}{u\ln(u)} + o(1) \ = \ \frac{\ln\ln(x)}{u(\ln\ln(x) - \ln\ln(y))} + o(1)$$

Since $u \geq 3$, we have $y = x^{1/u} \leq x^{1/3}$. Hence,

$$\frac{\ln\ln(x)}{u(\ln\ln(x)-\ln\ln(y))}+o(1) \ \le \ \frac{1}{u(1-\ln\ln(x^{1/3})/\ln\ln(x))}+o(1).$$

As $u \to \infty$ we know that $x \to \infty$ as well, hence the right hand side of the above inequality tends to 0. Thus, $\ln\ln(y)/(u\ln(u)) = o(1)$ as $u \to \infty$. For the next term, we note that,

$$\frac{(\ln(x))^2}{yu\ln(u)} \ = \ \frac{(\ln(x))^2}{x^{1/u}u\ln(u)} \ = \ \exp\left(\ln(2)+\ln\ln(x)-\frac{\ln(x)}{u}-\ln(u)-\ln\ln(u)\right).$$

Since $u \le (1-\epsilon)\ln(x)/\ln\ln(x) < \ln(x)/\ln\ln(x)$, we have

$$\begin{aligned}
&\exp\left(\ln(2)+\ln\ln(x)-\frac{\ln(x)}{u}-\ln(u)-\ln\ln(u)\right)\\
&\qquad < \ \exp\left(\ln(2)+\ln\ln(x)-\frac{\ln\ln(x)\ln(x)}{\ln(x)}-\ln(u)-\ln\ln(u)\right)\\
&\qquad\qquad = \ \exp\left(\ln(2)-\ln(u)-\ln\ln(u)\right).
\end{aligned}$$

Since $\exp\left(\ln(2)-\ln(u)-\ln\ln(u)\right) \to 0$ as $u \to \infty$, we have $(\ln(x))^2/(yu\ln(u)) = o(1)$ as $u \to \infty$.

Next we consider the term $R/(u\ln(u))$. Recall that,

$$R = \int_1^{\ln(y)} e^{s(\alpha/\ln(y))}V(e^s)\,ds$$

where $\alpha = \ln(u)+\ln\ln(u+1)$ and $V$ is a function connected with the error term in the Prime Ideal Theorem. Then,

$$\frac{R}{u\ln(u)} \ = \ \frac{1}{u\ln(u)}\int_1^{\ln(y)} e^{s(\alpha/\ln(y))}V(e^s)\,ds.$$

Performing the substitution $t = e^s$ in the above integral, we get

$$\begin{aligned}
&\frac{1}{u\ln(u)}\int_1^{\ln(y)} e^{s(\alpha/\ln(y))}V(e^s)\,ds \ = \ \frac{1}{u\ln(u)}\int_1^{\ln(y)} (e^s)^{\alpha/\ln(y)-1}\,V(e^s)e^s\,ds\\
&= \ \frac{1}{u\ln(u)}\int_e^y t^{\alpha/\ln(y)-1}V(t)\,dt \ \le \ \frac{y^{\alpha/\ln(y)-1}}{u\ln(u)}\int_e^y V(t)\,dt \ = \ \frac{e^{\alpha-\ln(y)}}{u\ln(u)}\int_e^y V(t)\,dt\\
&\qquad = \ \frac{\ln(u+1)}{y\ln(u)}\int_e^y V(t)\,dt.
\end{aligned}$$

Recalling that,

$$\int_e^y V(t)\,dt \;=\; O\left(ye^{-c\sqrt{\ln(y)}}\right)$$

where $c > 0$ is a constant depending on $K$ only, we get

$$\frac{\ln(u+1)}{y\ln(u)}\int_e^y V(t)\,dt \;=\; \frac{\ln(u+1)}{y\ln(u)}O\left(ye^{-c\sqrt{\ln(y)}}\right) \;=\; O\left(\frac{\ln(u+1)}{\ln(u)}e^{-c\sqrt{\ln\ln(x)}}\right)$$

with the last equality following from the fact that $y = x^{1/u}$ and $u < \ln(x)/\ln\ln(x)$. As mentioned above, since $u \leq (1-\epsilon)\ln(x)/\ln\ln(x)$, as $u \to \infty$ we know $x \to \infty$ as well. Hence, as $u \to \infty$, we have $(\ln(u+1))/(\ln(u)) \to 1$ and $e^{-c\sqrt{\ln\ln(x)}} \to 0$. Thus, $R/(u\ln(u)) = o(1)$. Hence, we have shown that

$$\exp\left(\frac{\ln(u+1)}{2} + O(\ln\ln(y)) + O\left(\frac{(\ln(x))^2}{y}\right) + O(R)\right) \;=\; u^{-u(o(1))}$$

which implies,

$$\begin{aligned}
&\Psi_K(x, x^{1/u})\\
&\leq\; \frac{x}{u^{u(1+o(1))}}u^{-u\left(-\frac{\ln(u+1)}{u\ln(u)} + O\left(\frac{\ln\ln(y)}{u\ln(u)}\right) + O\left(\frac{(\ln(x))^2}{yu\ln(u)}\right) + O\left(\frac{R}{u\ln(u)}\right)\right)}\\
&\qquad\qquad =\; \frac{x}{u^{u(1+o(1))}}u^{-u(o(1))} \;=\; \frac{x}{u^{u(1+o(1))}}.
\end{aligned}$$

Combining this with the inequality derived from Theorem 2.2.1, we have

$$\frac{x}{u^{u(1+o(1))}} \;\leq\; \Psi_K(x, x^{1/u}) \;\leq\; \frac{x}{u^{u(1+o(1))}},$$

which implies,

$$\Psi_K(x, x^{1/u}) \;=\; \frac{x}{u^{u(1+o(1))}}.$$

In order to replace the $x$ in the numerator of the above with $|S(x)|$, we note that $|S(x)|$ is proportional to $x$ asymptotically. Specifically, M.R. Murty and J. Van Order [9, Thm. 5] show that

$$|S(x)| \;=\; c_K\, x(1+o(1)), \tag{3.20}$$

where $c_K > 0$ is a constant dependent only on $K$. Noting that,

$$\frac{-\ln(c_K(1+o(1)))}{u\ln(u)} = o(1),$$

we have

$$c_K(1+o(1)) = \frac{1}{u^{o(1)u}}.$$

Hence,

$$\frac{|S(x)|}{u^{(1+o(1))u}} \;=\; \frac{c_K x(1+o(1))}{u^{(1+o(1))u}} \;=\; \left(\frac{1}{u^{o(1)u}}\right)\frac{x}{u^{(1+o(1))u}} \;=\; \frac{x}{u^{(1+o(1))u}}.$$

■

The utility of Corollary 3.1.3 stems from the following,

**Corollary 3.1.4.** *For $a > 0$,*

$$\Psi_K(x, L^a) \;=\; \frac{|S(x)|}{L^{1/2a+o(1)}}$$

*as $x \to \infty$.*

Proof. With $y = L^a$, we have

$$u \;=\; \frac{\ln(x)}{\ln(y)} \;=\; \frac{\ln(x)}{a\ln(L)} \;=\; \frac{\ln(x)}{a\sqrt{\ln(x)\ln\ln(x)}} \;=\; \frac{\sqrt{\ln(x)}}{a\sqrt{\ln\ln(x)}}.$$

By Corollary 3.1.3 we have,

$$\begin{aligned}
\Psi_K(x, L^a) &= |S(x)| \frac{\sqrt{\ln(x)}}{a\sqrt{\ln\ln(x)}}^{-\frac{\sqrt{\ln(x)}}{a\sqrt{\ln\ln(x)}}(1+o(1))} \\
&= |S(x)| \exp\left(-\ln\left(\frac{\sqrt{\ln(x)}}{a\sqrt{\ln\ln(x)}}\right)\left(\frac{\sqrt{\ln(x)}}{a\sqrt{\ln\ln(x)}}\right)(1+o(1))\right) \\
&= |S(x)| \exp\left(-\left(\frac{1}{2}\ln\ln(x) - \ln(a) - \frac{1}{2}\ln\ln\ln(x)\right)\left(\frac{\sqrt{\ln(x)}}{a\sqrt{\ln\ln(x)}}\right)(1+o(1))\right) \\
&= |S(x)| \exp\left(-\sqrt{\ln(x)\ln\ln(x)}\left(\frac{1}{2a} - \frac{\ln(a)}{a\ln\ln(x)} - \frac{\ln\ln\ln(x)}{2a\ln\ln(x)}\right)(1+o(1))\right) \\
&= |S(x)| \exp\left(-\sqrt{\ln(x)\ln\ln(x)}\left(\frac{1}{2a} + o(1)\right)\right) \;=\; \frac{|S(x)|}{L^{1/2a+o(1)}}.
\end{aligned}$$

as $x \to \infty$.

■

Note that in Corollary 3.1.4 we are considering the $o(1)$ term with respect to $x$ rather than $u$. This is due to the fact that $u = \sqrt{\ln(x)}/(a\sqrt{\ln\ln(x)})$ is a function of $x$ alone. Corollary 3.1.4 will be crucial in our proof of the upper bound. Let us restate our claim explicitly,

**Theorem 1(a).** (The Upper Bound) *Let $K$ be a degree $n$ extension of $\mathbb{Q}$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. Let $\Omega$ be the smallest integer such that the first $\Omega$ terms of an ideal sequence generated by randomly and independently selecting ideals with uniform probability from the set $S(x)$ contains a square dependent subsequence. Then for any $\epsilon > 0$ the probability that $\Omega < L^{\sqrt{2}+\epsilon}$ tends to 1 as $x \to \infty$.*

Proof. Let $\epsilon > 0$ and $x \geq 2$ be given. Suppose we were to generate an ideal sequence of length $[L^{\sqrt{2}+\epsilon}]$ by choosing ideals from $S(x)$ independently and with uniform probability. For any positive integer $k$ and $y \geq 2$, the probability that $k$ ideals from our sequence are $y$-smooth is,

$$\binom{[L^{\sqrt{2}+\epsilon}]}{k}\left(\frac{\Psi_K(x,y)}{|S(x)|}\right)^k\left(1-\frac{\Psi_K(x,y)}{|S(x)|}\right)^{[L^{\sqrt{2}+\epsilon}]-k}$$

We want to show that there exists a smoothness bound $y$, such that the probability we have fewer than $\pi_K(y)+1$, $y$-smooth ideals in our sequence, tends to 0 as $x \to \infty$. Let $y = L^{1/\sqrt{2}}$. By Corollary 3.1.4 we have,

$$\frac{\Psi_K(x, L^{1/\sqrt{2}})}{|S(x)|} \;=\; \frac{1}{L^{1/\sqrt{2}+o(1)}}$$

as $x \to \infty$. Hence, the probability that at most $\pi_K(L^{1/\sqrt{2}})$ ideals from our sequence are $y$-smooth is

$$\sum_{k=1}^{\pi_K(L^{1/\sqrt{2}})}\binom{[L^{\sqrt{2}+\epsilon}]}{k}\left(\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^k\left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{[L^{\sqrt{2}+\epsilon}]-k}.$$

By the Prime Ideal Theorem we have $\pi_K(L^{1/\sqrt{2}}) \leq \sqrt{2}(1+\delta)L^{1/\sqrt{2}}/\ln(L)$ for sufficiently large $x$, where $\delta > 0$ is arbitrary. For ease of notation let $N = [\sqrt{2}(1+\delta)L^{1/\sqrt{2}}/\ln(L)]$. Then the above sum is bounded above by,

$$\sum_{k=1}^{N}\binom{[L^{\sqrt{2}+\epsilon}]}{k}\left(\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^k\left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{[L^{\sqrt{2}+\epsilon}]-k}$$

$$=\;\sum_{k=1}^{N}\binom{[L^{\sqrt{2}+\epsilon}]}{k}\left(\frac{1}{L^{1/\sqrt{2}+o(1)}-1}\right)^k\left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{L^{\sqrt{2}+\epsilon}},$$

for sufficiently large $x$, with the last equality following from a rearrangement of terms and the fact that we can write $L^{\sqrt{2}+\epsilon} = [L^{\sqrt{2}+\epsilon}] + \Delta$, where $0 \leq \Delta < 1$, and

$$\lim_{x\to\infty}\left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{\Delta} \;=\; 1.$$

That is, considering $L^{\sqrt{2}+\epsilon}$ rather than $[L^{\sqrt{2}+\epsilon}]$ will not affect the limit. For $1 \leq k \leq n$ we have the well known inequality $\binom{n}{k} < (ne/k)^k$. Therefore,

$$\sum_{k=1}^{N} \binom{[L^{\sqrt{2}+\epsilon}]}{k} \left(\frac{1}{L^{1/\sqrt{2}+o(1)}-1}\right)^k \left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{L^{\sqrt{2}+\epsilon}}$$

$$< \sum_{k=1}^{N} \left(\frac{eL^{\sqrt{2}+\epsilon}}{kL^{1/\sqrt{2}+o(1)}}\right)^k \left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{L^{\sqrt{2}+\epsilon}}$$

$$= \sum_{k=1}^{N} \left(\frac{eL^{1/\sqrt{2}+\epsilon+o(1)}}{k}\right)^k \left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{L^{\sqrt{2}+\epsilon}}$$

Considering the function $(y/k)^k$ with respect to $k$, we have

$$\frac{d}{dk}\left(\frac{y}{k}\right)^k = (\ln(y)-\ln(k)-1)\left(\frac{y}{k}\right)^k.$$

We can see that $(y/k)^k$ is increasing for $k \in (0, y/e)$ and has a maximum at $k = y/e$. Since $N < L^{1/\sqrt{2}+\epsilon+o(1)}$, we see that $\left(eL^{1/\sqrt{2}+\epsilon+o(1)}/k\right)^k$ is maximized, with respect to the $k$ in our sum, when $k = N$. Hence,

$$\sum_{k=1}^{N} \left(\frac{eL^{1/\sqrt{2}+\epsilon+o(1)}}{k}\right)^k \left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{L^{\sqrt{2}+\epsilon}}$$

$$\leq \sum_{k=1}^{N} \left(\frac{eL^{1/\sqrt{2}+\epsilon+o(1)}}{N}\right)^N \left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{L^{\sqrt{2}+\epsilon}}$$

$$= N\left(\frac{eL^{1/\sqrt{2}+\epsilon+o(1)}}{N}\right)^N \left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{L^{\sqrt{2}+\epsilon}}.$$

Noting that

$$\frac{eL^{1/\sqrt{2}+\epsilon+o(1)}}{N} = \frac{eL^{\epsilon+o(1)}\ln(L)}{\sqrt{2}(1+\delta)} = \left(\frac{e\ln(L)}{\sqrt{2}(1+\delta)}\right)L^{\epsilon+o(1)} = L^{\epsilon+o(1)},$$

and,

$$N = \frac{\sqrt{2}(1+\delta)L^{1/\sqrt{2}}}{\ln(L)} = \left(\frac{\sqrt{2}(1+\delta)}{\ln(L)}\right)L^{1/\sqrt{2}} = L^{1/\sqrt{2}+o(1)},$$

we have

$$N\left(\frac{eL^{1/\sqrt{2}+\epsilon+o(1)}}{N}\right)^{N}\left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)^{L^{\sqrt{2}+\epsilon}}$$

$$= \exp\left(\left(\frac{1}{\sqrt{2}}+o(1)\right)\ln(L)+L^{1/\sqrt{2}+o(1)}(\epsilon+o(1))\ln(L)+L^{\sqrt{2}+\epsilon}\ln\left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)\right)$$

$$= \exp\left(\left(\frac{1}{\sqrt{2}}+o(1)\right)\ln(L)+L^{1/\sqrt{2}+o(1)}+L^{\sqrt{2}+\epsilon}\ln\left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)\right).$$

Recalling that for $|x|<1$ we can write $\ln(1-x)=-\sum_{\ell=1}^{\infty}x^{\ell}/\ell$, we have

$$\exp\left(\left(\frac{1}{\sqrt{2}}+o(1)\right)\ln(L)+L^{1/\sqrt{2}+o(1)}+L^{\sqrt{2}+\epsilon}\ln\left(1-\frac{1}{L^{1/\sqrt{2}+o(1)}}\right)\right)$$

$$= \exp\left(\left(\frac{1}{\sqrt{2}}+o(1)\right)\ln(L)+L^{1/\sqrt{2}+o(1)}-L^{\sqrt{2}+\epsilon}\sum_{\ell=1}^{\infty}\frac{1}{\ell L^{\ell/\sqrt{2}+o(1)}}\right)$$

$$= \exp\left(\left(\frac{1}{\sqrt{2}}+o(1)\right)\ln(L)+L^{1/\sqrt{2}+o(1)}-L^{1/\sqrt{2}+\epsilon+o(1)}-L^{\epsilon+o(1)}+o(1)\right).$$

Noting that

$$\left(\frac{1}{\sqrt{2}}+o(1)\right)\ln(L)+L^{1/\sqrt{2}+o(1)}-L^{1/\sqrt{2}+\epsilon+o(1)}-L^{\epsilon+o(1)}+o(1) \ \to\ -\infty,$$

as $x\to\infty$, we have

$$\lim_{x\to\infty}\exp\left(\left(\frac{1}{\sqrt{2}}+o(1)\right)\ln(L)+L^{1/\sqrt{2}+o(1)}-L^{1/\sqrt{2}+\epsilon+o(1)}-L^{\epsilon+o(1)}+o(1)\right) \ = \ 0.$$

That is, the probability that we have fewer than $\pi_K(L^{1/\sqrt{2}})+1$ ideals in our sequence that are $y$-smooth, tends to 0 as $x\to\infty$. Therefore, the probability that we have at least $\pi_K(L^{1/\sqrt{2}})+1$ ideals in our sequence that are $y$-smooth, tends to 1 as $x\to\infty$. This implies that the probability our sequence has a square dependent subsequence also tends to 1 as $x\to\infty$, by the argument from linear algebra given in Chapter 2.

■

## 3.2 The Lower Bound

We will now show the probability that $\Omega \leq L^{\sqrt{2}-\epsilon}$ tends to 0 as $x \to \infty$. The argument is a generalization of that given by Pomerance [13, Thm. 1 (lower bound)] when $K = \mathbb{Q}$. We begin with some preliminary propositions.

**Proposition 1.** *For a given $x \geq 2$, let $A_1, \ldots, A_k$ be subsets of ideals from the set $S(x)$. Consider a random sequence of ideals from $S(x)$ of length at most $\ell$, where the ideals in the sequence need not be distinct. The probability that this sequence contains $k$ distinct terms $I_1, I_2, \ldots, I_k$ with $I_1 \in A_1$, $I_2 \in A_2$, $\ldots$, $I_k \in A_k$ is at most*

$$\frac{\ell^k}{|S|^k}|A_1|\cdots|A_k|$$

Proof. For $i \in \{1, 2, \ldots, k\}$, the probability that a given ideal of $S(x)$ is in $A_i$ is

$$\frac{|A_i|}{|S(x)|}.$$

Therefore, the probability that any of the ideals in our sequence of length $\ell$ are in $A_i$ is,

$$\frac{\ell|A_i|}{|S(x)|}.$$

Hence, the probability that each of the $A_i$ contains an ideal from our sequence is *at most* the product of above probabilities for each $A_i$,

$$\prod_{i=1}^{k} \frac{\ell}{|S(x)|}|A_i| \;=\; \frac{\ell^k}{|S(x)|^k}|A_1|\cdots|A_k|.$$

■

**Proposition 2.** *Consider a random ideal sequence drawn from $S(x)$ of length at most $L^{1.5}$. The probability it has a term divisible by the square of a prime ideal $\mathfrak{p}$ with $N(\mathfrak{p}) > L^3$ tends to 0 as $x \to \infty$.*

Proof. Note that if the square of a prime ideal $\mathfrak{p}$ divides an ideal $I$ with $N(I) \leq x$, it must be the case that $N(\mathfrak{p}) \leq \sqrt{x}$. Let $\mathfrak{p}$ be prime ideal with $L^3 < N(\mathfrak{p}) \leq \sqrt{x}$. We will employ Proposition 1 with $k = 1$ and $A_1 = \{I \subseteq \mathcal{O}_K : N(I) \leq x;\, \mathfrak{p}^2 \mid I\}$. Doing so we get the probability that our random ideal sequence of length $L^{1.5}$ has a term divisible by $\mathfrak{p}^2$ is at most,

$$\frac{L^{1.5}}{|S(x)|}|\{I \subseteq \mathcal{O}_K : N(I) \leq x;\, \mathfrak{p}^2 \mid I\}|.$$

Summing over all prime ideals $\mathfrak{p}$ with $L^3 < N(\mathfrak{p}) \leq \sqrt{x}$, we see that our desired probability is at most,

$$\sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K \\ L^3<N(\mathfrak{p})\leq\sqrt{x}}} \frac{L^{1.5}\cdot|\{I \subseteq \mathcal{O}_K : N(I) \leq x;\, \mathfrak{p}^2 \mid I\}|}{|S(x)|} = \sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K \\ L^3<N(\mathfrak{p})\leq\sqrt{x}}} \frac{L^{1.5}\cdot|S(x/N(\mathfrak{p})^2)|}{|S(x)|}$$

Consider the set of ideals $\{IJ : I \in S(x/N(\mathfrak{p})^2);\, J \in S(N(\mathfrak{p})^2)\}$, where $\mathfrak{p}$ is a prime ideal satisfying $L^3 < N(\mathfrak{p})^2 \leq \sqrt{x}$. Since,

$$\{IJ : I \in S(x/N(\mathfrak{p})^2);\, J \in S(N(\mathfrak{p})^2)\} \;\subseteq\; S(x)$$

we know,

$$|\{IJ : I \in S(x/N(\mathfrak{p})^2);\, J \in S(N(\mathfrak{p})^2)\}| \;\leq\; |S(x)|.$$

Furthermore,

$$|\{IJ : I \in S(x/N(\mathfrak{p})^2);\, J \in S(N(\mathfrak{p})^2)\}| \;=\; |S(x/N(\mathfrak{p})^2)||S(N(\mathfrak{p})^2)|,$$

which gives us,

$$\sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K\\ L^3<N(\mathfrak{p})\leq\sqrt{x}}} \frac{L^{1.5}\cdot|S(x/N(\mathfrak{p})^2)|}{|S(x)|} \;\leq\; \sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K\\ L^3<N(\mathfrak{p})\leq\sqrt{x}}} \frac{L^{1.5}\cdot|S(x/N(\mathfrak{p})^2)|}{|S(x/N(\mathfrak{p})^2)||S(N(\mathfrak{p})^2)|}$$

$$= \sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K\\ L^3<N(\mathfrak{p})\leq\sqrt{x}}} \frac{L^{1.5}}{|S(N(\mathfrak{p})^2)|} \;\leq\; \sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K\\ L^3<N(\mathfrak{p})\leq\sqrt{x}}} \frac{L^{1.5}}{(1-o(1))N(\mathfrak{p})^2},$$

with the last inequality following from (3.20). Simplifying further, we have

$$\sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K\\ L^3<N(\mathfrak{p})\leq\sqrt{x}}} \frac{L^{1.5}}{(1-o(1))N(\mathfrak{p})^2} \;\leq\; \left(\frac{n}{(1-o(1))L^{1.5}}\right)\sum_{L^3<p\leq\sqrt{x}}\frac{1}{p}.$$

Since $\sum_{p\leq y} 1/p = \ln\ln(y) + B + o(1)$, with $B \approx 0.2615$ being Mertens constant [3, Thm. 1.4.2], we have

$$\left(\frac{n}{(1-o(1))L^{1.5}}\right)\sum_{L^3<p\leq\sqrt{x}}\frac{1}{p} \;\leq\; \left(\frac{n}{(1-o(1))L^{1.5}}\right)\big(\ln\ln(\sqrt{x}) + B + o(1)\big)$$

Finally we note that,

$$\lim_{x\to\infty}\left(\frac{n}{(1-o(1))L^{1.5}}\right)\big(\ln\ln(\sqrt{x}) + B + o(1)\big) \;=\; 0.$$

■

**Proposition 3.** *Consider a random ideal sequence drawn from $S(x)$ of length at most $L^{1.5}$. The probability it has two terms divisible by a prime ideal $\mathfrak{p}$ with $N(\mathfrak{p}) > L^3$ tends to 0 as $x \to \infty$*

Proof. Let $\mathfrak{p}$ be a prime ideal with $L^3 < N(\mathfrak{p}) \leq x$. By Proposition 1, the probability that our random ideal sequence of length $L^{1.5}$ has two terms divisible by $\mathfrak{p}$ is at most,

$$\frac{L^3}{|S(x)|^2}|\{I \subseteq \mathcal{O}_K : N(I) \leq x;\ \mathfrak{p} \mid I\}|^2.$$

Summing over all prime ideals $\mathfrak{p}$ with $L^3 < N(\mathfrak{p}) \leq x$, we see that our desired probability is at most,

$$\sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K\\ L^3<N(\mathfrak{p})\le x}} \frac{L^3}{|S(x)|^2}|\{I\subseteq\mathcal{O}_K : N(I)\le x;\ \mathfrak{p}\mid I\}|^2 \;=\; \sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K\\ L^3<N(\mathfrak{p})\le x}} \frac{L^3|S(x/N(\mathfrak{p}))|^2}{|S(x)|^2}.$$

By (3.20) and the argument given in the proof Proposition 2 we have,

$$\sum_{\substack{\mathfrak{p}\subseteq\mathcal{O}_K\\ L^3<N(\mathfrak{p})\le x}} \frac{L^3|S(x/N(\mathfrak{p}))|^2}{|S(x)|^2} \;\le\; \frac{nL^3}{(1-o(1))}\sum_{L^3<p\le x}\frac{1}{p^2}.$$

Note that,

$$\frac{nL^3}{(1-o(1))}\sum_{L^3<p\le x}\frac{1}{p^2} \;=\; \frac{nL^3}{(1-o(1))}\sum_{\substack{m\in\mathbb{N}\\ L^3<m\le x}}\frac{1}{m^2}(\pi(m)-\pi(m-1))$$

Since $1/x^2$ is continuous and $\pi(x)$ is monotonically increasing on the interval $[L^3, x]$, we can write the above sum in terms of a Reimann-Stieltjes integral,

$$\frac{nL^3}{(1-o(1))}\sum_{\substack{m\in\mathbb{N}\\ L^3<m\le x}}\frac{1}{m^2}(\pi(m)-\pi(m-1)) \;=\; \frac{nL^3}{(1-o(1))}\int_{L^3}^{x}\frac{1}{t^2}d\pi(t).$$

Performing integration by parts we get,

$$\frac{nL^3}{(1-o(1))}\int_{L^3}^{x}\frac{1}{t^2}d\pi(t) \;=\; \frac{nL^3}{(1-o(1))}\left(\frac{\pi(x)}{x^2}-\frac{\pi(L^3)}{L^6}+2\int_{L^3}^{x}\frac{\pi(t)}{t^3}\,dt\right).$$

By a theorem of Chebyshev [3, Thm. 1.1.3] we know that for $t\ge 3$ there exists a positive number $B$ such that $\pi(t) < Bt/\ln(t)$. Hence,

$$\begin{aligned}
&\frac{nL^3}{(1-o(1))}\left(\frac{\pi(x)}{x^2}-\frac{\pi(L^3)}{L^6}+2\int_{L^3}^{x}\frac{\pi(t)}{t^3}\,dt\right)\\
&\qquad < \frac{nL^3}{(1-o(1))}\left(\frac{\pi(x)}{x^2}-\frac{\pi(L^3)}{L^6}+2\int_{L^3}^{x}\frac{B}{t^2\ln(t)}\,dt\right)\\
&\qquad \le \frac{nL^3}{(1-o(1))}\left(\frac{\pi(x)}{x^2}-\frac{\pi(L^3)}{L^6}+\frac{2B}{3\ln(L)}\int_{L^3}^{x}\frac{1}{t^2}\,dt\right)\\
&\qquad = \frac{nL^3}{(1-o(1))}\left(\frac{\pi(x)}{x^2}-\frac{\pi(L^3)}{L^6}+\frac{2B}{3\ln(L)L^3}-\frac{2B}{3x\ln(L)}\right)\\
&\qquad = \frac{n}{(1-o(1))}\left(\frac{\pi(x)L^3}{x^2}-\frac{\pi(L^3)}{L^3}+\frac{2B}{3\ln(L)}-\frac{2BL^3}{3x\ln(L)}\right).
\end{aligned}$$

The result follows by noting that

$$\lim_{x\to\infty} n(1+o(1))\left(\frac{\pi(x)L^3}{x^2} - \frac{\pi(L^3)}{L^3} + \frac{2B}{3\ln(L)} - \frac{2BL^3}{3x\ln(L)}\right) \;=\; 0.$$

■

Suppose a given ideal $I \subseteq \mathcal{O}_K$ has prime ideal factorization $I = \mathfrak{p}_1\mathfrak{p}_2\cdots\mathfrak{p}_\ell$, where $N(\mathfrak{p}_1) \geq N(\mathfrak{p}_2) \geq \cdots \geq N(\mathfrak{p}_\ell)$. Note that these prime ideals are not necessarily distinct, and that $N(\mathfrak{p}_i) = N(\mathfrak{p}_j)$ does not imply $\mathfrak{p}_i = \mathfrak{p}_j$. Given an ordering of the prime ideals in the factorization of $I$ by norm, let $P_k(I) = \mathfrak{p}_k$, with $P_k(I) = \mathcal{O}_k$ for $k > \ell$.

**Proposition 4.** *For any fixed positive integer $k$, the number of $L^3$-smooth ideals $I \subseteq \mathcal{O}_K$ with $N(I) \leq x$, for which $N(P_k(I)) \leq L^{1/3}$ is at most $|S(x)|/L^{1.5+o(1)}$.*

Proof. Let $\mathcal{B}$ be the set of ideals in $\mathcal{O}_K$ with norm at most $x$, whose prime ideal factorization consists of at most $k-1$ prime ideals, each with norm in the interval $(L^{1/3}, L^3]$. Then the ideals that we wish to count, namely, the ideals $I \in S(x)$ such that $N(P_k(I)) \leq L^{1/3}$, are of the form $JB$, where $J \subseteq \mathcal{O}_K$ is $L^{1/3}$-smooth and $B \in \mathcal{B}$. The number of such ideals is exactly,

$$\sum_{B\in\mathcal{B}} \Psi_K\left(\frac{x}{N(B)}, L^{1/3}\right).$$

By Corollary 3.1.2 and (3.20), we have

$$\Psi_K\left(\frac{x}{N(B)}, L^{1/3}\right) \;=\; \frac{S(x/N(B))}{L^{1.5+o(1)}} \;\leq\; \frac{c_K(1+o(1))x}{N(B)\,L^{1.5+o(1)}} \;=\; \frac{x}{N(B)\,L^{1.5+o(1)}},$$

with the last equality following from the fact that $c_K(1+o(1)) = 1/L^{o(1)}$. However, by (3.20) we have

$$\frac{|S(x)|}{N(B)L^{1.5+o(1)}} \;\leq\; \frac{x(1+o(1))}{N(B)L^{1.5+o(1)}} \;=\; \frac{x}{N(B)L^{1.5+o(1)}} \;\leq\; \frac{|S(x)|}{N(B)L^{1.5+o(1)}}$$

Thus,

$$\sum_{B\in\mathcal{B}} \Psi_K\left(\frac{x}{N(B)}, L^{1/3}\right) \;=\; \sum_{B\in\mathcal{B}} \frac{|S(x)|}{N(B)\,L^{1.5+o(1)}} \;=\; \frac{|S(x)|}{L^{1.5+o(1)}}\sum_{B\in\mathcal{B}}\frac{1}{N(B)}$$

Considering just the sum $\sum_{B\in\mathcal{B}} 1/N(B)$ for a moment, we see that

$$0 \;<\; \sum_{B\in\mathcal{B}}\frac{1}{N(B)} \;\le\; \left(1+\sum_{L^{1/3}<N(\mathfrak{p})\le L^3}\frac{1}{N(\mathfrak{p})}\right)^{k-1} \;\le\; \left(1+n\sum_{L^{1/3}<p\le L^3}\frac{1}{p}\right)^{k-1}.$$

Recalling that $\sum_{p\le x} 1/p = \ln\ln(x) + B + o(1)$, where $B \approx 0.2615$ [3, Thm. 1.4.2]. Hence,

$$\left(1+n\sum_{L^{1/3}<p\le L^3}\frac{1}{p}\right)^{k-1} \;=\; \left(1+n\Big(\ln\ln(L^3)-\ln\ln(L^{1/3})+o(1)\Big)\right)^{k-1}$$

$$=\; (1+2n\ln(3)+o(1))^{k-1}$$

Since,

$$\frac{-(k-1)\ln(1+2n\ln(3)+o(1))}{\ln(L)} \;=\; o(1)$$

we have,

$$(1+2n\ln(3)+o(1))^{k-1} \;=\; \frac{1}{L^{o(1)}},$$

which implies,

$$\sum_{B\in\mathcal{B}}\frac{1}{N(B)} \;\le\; \frac{1}{L^{o(1)}}.$$

Hence,

$$\frac{|S(x)|}{L^{1.5+o(1)}}\sum_{B\in\mathcal{B}}\frac{1}{N(B)} \;\le\; \frac{|S(x)|}{L^{1.5+o(1)}}.$$

■

**Proposition 5.** *For any fixed positive integer $k$, we have for $a \in [1/3, 3]$ and $I \in S(L^{3k})$, that the number of ideals $J \in S(x)$, having the properties $I \mid J$, $J$ is $L^3$-smooth, and $N(P_k(J)) \leq L^a$ is at most $|S(x)|/(N(I)\, L^{1/(2a)+o(1)})$.*

Proof. The following proof is very similar to that of the previous proposition. Let $a \in [1/3, 3]$, $k \in \mathbb{N}$, and $I \in S(L^{3k})$. Note that if $N(P_k(I)) > L^a$, the set of ideals $J$ satisfying the properties specified in the Proposition would be the empty set, since $I \mid J$ implies $N(P_k(J)) > L^a$. Therefore, we may assume that $N(P_k(I)) \leq L^a$. Let $\mathcal{B}$ be the set of ideals in $\mathcal{O}_K$ with norm at most $x$, that are composed of at most $k-1$ prime ideals, each with norm in the interval $(L^a, L^3]$. Then the ideals that we wish to count are of the form $IHB$, where $H \subseteq \mathcal{O}_K$ is $L^a$-smooth and $B \in \mathcal{B}$. The number of such ideals is exactly,

$$\sum_{B \in \mathcal{B}} \Psi_K\left(\frac{x}{N(I)N(B)}, L^a\right).$$

By Corollary 3.1.3 and (3.20), we the can make the same argument as that given in the proof of Proposition 4 to show

$$\Psi_K\left(\frac{x}{N(B)N(I)}, L^a\right) = \frac{|S(x)|}{N(B)N(I)\, L^{1/(2a)\, +\, o(1)}}.$$

Thus,

$$\begin{aligned}\sum_{B \in \mathcal{B}} \Psi_K\left(\frac{x}{N(B)N(I)}, L^a\right) &= \sum_{B \in \mathcal{B}} \frac{|S(x)|}{N(B)N(I)\, L^{1/(2a)\, +\, o(1)}} \\ &= \frac{|S(x)|}{N(I) L^{1/(2a)\, +\, o(1)}} \sum_{B \in \mathcal{B}} \frac{1}{N(B)}.\end{aligned}$$

As in the proof of Proposition 4, we will consider just the sum $\sum_{B \in \mathcal{B}} 1/N(B)$ for a moment, and note that

$$0 \ \le \ \sum_{B\in\mathcal{B}} \frac{1}{N(B)} \ \le \ \left(1 + \sum_{L^a<N(\mathfrak{p})\le L^3} \frac{1}{N(\mathfrak{p})}\right)^{k-1}$$

$$\le \ \left(1+n\sum_{L^a<p\le L^3}\frac{1}{p}\right)^{k-1} \ = \ \left(1 - n\ln(3)+n\ln(a)+o(1)\right)^{k-1}.$$

Since, $(1-n\ln(3)+n\ln(a)+o(1))^{k-1} = 1/L^{o(1)}$, we have

$$= \ \frac{|S(x)|}{N(I)L^{1/(2a)+o(1)}}\sum_{B\in\mathcal{B}}\frac{1}{N(B)} \ \le \ \frac{|S(x)|}{N(I)L^{1/(2a)+o(1)}}.$$

■

**Proposition 6.** *For any fixed positive integer $k$, the number of $L^3$-smooth ideals $I \in S(x)$ such that $P_k(I)^2 \mid I$, is at most $|S(x)|/L^{\sqrt{2}+o(1)}$ as $x \to \infty$.*

Proof. Let $k \in \mathbb{N}$. We begin by counting only those ideals $L^3$-smooth ideals $I \in S(x)$ such that $1/3 < N(P_k(I)) \le L^3$ and $P_k(I)^2 \mid I$. To do so, we partition the interval $(L^{1/3}, L^3]$ into disjoint intervals of the form $(L^{a_{i-1}}, L^{a_i}]$, so as to employ Proposition 5. To do so, let $T(i) = (e^{i-1}L^{1/3}, e^iL^{1/3}]$, for $i = 1, 2, \ldots, \lceil(8/3)\ln(L)\rceil$. Note that $T(i) \cap T(j) = \emptyset$ for $i \ne j$, and

$$\bigcup_{i=1}^{\lceil(8/3)\ln(L)\rceil} T(i) \ = \ \bigcup_{i=1}^{\lceil(8/3)\ln(L)\rceil} (e^{i-1}L^{1/3}, e^iL^{1/3}] \ \supseteq \ (L^{1/3}, L^3].$$

Let $a_i = 1/3 + i/\ln(L)$. Note that,

$$L^{a_i} = L^{1/3+i/\ln(L)} \ = \ L^{1/3}e^{(i/\ln(L))\ln(L)} \ = \ e^iL^{1/3}.$$

Hence, $T(i) = (L^{a_{i-1}}, L^{a_i}]$. Letting $L^{a_i} = L^3$ for $i = \lceil(8/3)\ln(L)\rceil$ we have created our desired partition.

Let $A(i)$ be the set of $L^3$-smooth ideals $I \in S(x)$ such that $P_k(I)^2 \mid I$ and $N(P_k(I)) \in T(i)$. If $I \subseteq A(i)$, then it must be the case that $I$ is $L^3$-smooth, $I \in S(x)$,

$N(P_k(I)) \leq L^{a_i}$, and there exists a prime ideal $\mathfrak{p} \subseteq \mathcal{O}_K$ with $N(\mathfrak{p}) \in T(i)$ such that $\mathfrak{p}^2 \mid I$. Note that this is a necessary condition for an ideal to be contained in $A(i)$, but it is not sufficient. Hence, the set of ideals that satisfy the aforementioned conditions actually contains the set $A(i)$, however, these conditions allow us to employ Proposition 5.

Let $1 \leq i \leq \lceil (8/3)\ln(L) \rceil$ and $\mathfrak{p}$ be a prime ideal such that $N(\mathfrak{p}) \in T(i)$. Since $k$ is a positive integer, $a_i \in [1/3, 3]$, and $\mathfrak{p}^2 \in S(L^{3k})$, by Proposition 5 the number of $L^3$-smooth ideals $I \in S(x)$, such that $N(P_k(I)) \leq L^{a_i}$ and $\mathfrak{p}^2 \mid I$ is at most $|S(x)|/(N(\mathfrak{p}^2)L^{1/(2a)+o(1)})$. By summing over all prime ideals $\mathfrak{p}$ with $N(\mathfrak{p}) \in T(i)$, we get

$$A(i) \;\leq \sum_{N(\mathfrak{p}) \in T(i)} \frac{|S(x)|}{N(\mathfrak{p}^2)L^{1/(2a_i)+o(1)}} \;=\; \frac{|S(x)|}{L^{1/(2a_i)+o(1)}} \sum_{N(\mathfrak{p}) \in T(i)} \frac{1}{N(\mathfrak{p}^2)}$$

Considering just the above sum for a moment, we see that

$$\sum_{N(\mathfrak{p}) \in T(i)} \frac{1}{N(\mathfrak{p}^2)} \;\leq\; \frac{1}{L^{2(a_{i-1})}} \sum_{N(\mathfrak{p}) \in T(i)} 1 \;\leq\; \frac{n(L^{a_i} - L^{a_{i-1}})}{L^{2(a_{i-1})}}$$

$$= \;\frac{n(L^{1/3}L^{i/\ln(L)} - L^{1/3}L^{(i-1)/\ln(L)})}{L^{2/3}L^{(2i-2)/\ln(L)}} \;=\; \left(\frac{1}{L^{1/3}L^{i/\ln(L)}}\right) \frac{n(1 - L^{-1/\ln(L)})}{L^{-2/\ln(L)}}$$

$$= \;\left(\frac{1}{L^{1/3}L^{i/\ln(L)}}\right) \cdot n\left(L^{2/\ln(L)} - L^{1\ln(L)}\right) \;=\; \left(\frac{1}{L^{a_i}}\right) \cdot n(e^2 - e) \tag{3.21}$$

Since $n(e^2 - e) = 1/L^{o(1)}$ we have,

$$A(i) \;\leq\; \left(\frac{|S(x)|}{L^{1/(2a_i)+o(1)}}\right)\left(\frac{1}{L^{a_i+o(1)}}\right) \;=\; \frac{|S(x)|}{L^{a_i+1/(2a_i)+o(1)}}$$

Next we note that $a + 1/(2a)$ has a minimum value of $\sqrt{2}$ when $a = 2/\sqrt{2}$. Hence, for $i = 1, 2, \ldots, \lceil (8/3)\ln(L) \rceil$, we have

$$A(i) \;\leq\; \frac{|S(x)|}{L^{\sqrt{2}+o(1)}}.$$

Therefore, the number of ideals $I \subseteq \mathcal{O}_K$ such that $N(I) \leq x$, $P_k(I)^2 \mid I$, and $N(P_k(I)) \in (L^{1/3}, L]$ is at most

$$\leq \lceil (8/3)\ln(L) \rceil \frac{|S(x)|}{L^{\sqrt{2}+o(1)}}.$$

Noting that $\lceil (8/3)\ln(L) \rceil = 1/L^{o(1)}$, we get

$$\lceil (8/3)\ln(L) \rceil \frac{|S(x)|}{L^{\sqrt{2}+o(1)}} \;=\; \left(\frac{1}{L^{o(1)}}\right) \frac{|S(x)|}{L^{\sqrt{2}+o(1)}} \;=\; \frac{|S(x)|}{L^{\sqrt{2}+o(1)}}.$$

Finally, we want to consider the number of ideals $I \in S(x)$ such that $P_k(I)^2 \mid I$ and $N(P_k(I)) \leq L^{1/3}$. Let $C$ denote this set of ideals. By Proposition 4 we have

$$|C| \;\leq\; \frac{|S(x)|}{L^{1.5+o(1)}}$$

Hence, the number of $L^3$-smooth ideals $I \in S(x)$ such that $P_k(I)^2 \mid I$, is at most

$$|C| + \frac{|S(x)|}{L^{\sqrt{2}+o(1)}} \;\leq\; \frac{|S(x)|}{L^{1.5+o(1)}} + \frac{|S(x)|}{L^{\sqrt{2}+o(1)}} \;=\; \frac{|S(x)|}{L^{\sqrt{2}+o(1)}}\left(1 + \frac{1}{L^{1.5-\sqrt{2}}}\right).$$

The proposition follows by taking $x \to \infty$.

■

We are now ready to prove the lower bound. Let us restate our claim explicitly,

**Theorem 1(b).** (The Lower Bound) *Let $K$ be a degree $n$ extension of $\mathbb{Q}$, and let $\mathcal{O}_K$ be the ring of algebraic integers in $K$. Let $\Omega$ be the smallest integer such that the first $\Omega$ terms of an ideal sequence generated by randomly and independently selecting ideals with uniform probability from the set $S(x)$ contains a square dependent subsequence. Then for any $\epsilon > 0$ the probability that $\Omega \leq L^{\sqrt{2}-\epsilon}$ tends to 0 as $x \to \infty$.*

Proof. Let $x \geq 2$, and let $\epsilon > 0$ be given. Suppose we have a sequence of ideals of length $[L^{\sqrt{2}-\epsilon}]$ drawn from $S(x)$ that has a square dependent subsequence. Call this subsequence $D$. We would like to determine the probability of such a subsequence existing, and show this probability tends to 0 as $x \to \infty$.

We begin by supposing there is an ideal $I$ in $D$ that is not $L^3$-smooth. Since $D$ is square dependent, there must exist a prime ideal $\mathfrak{p} \subseteq \mathcal{O}_K$ with $N(\mathfrak{p}) > L^3$ such that either $\mathfrak{p}^2 \mid I$ or $\mathfrak{p} \mid I$ and $\mathfrak{p} \mid J$ where $J$ is another ideal of $D$. However, by Propositions 2 and 3 we know the probability that $D$ has a term divisible by the square of a prime ideal $\mathfrak{p}$ with $N(\mathfrak{p}) > L^3$, or two terms each divisible by $\mathfrak{p}$ tends to 0 as $x \to \infty$. Hence, we may assume that each term of $D$ is $L^3$-smooth.

Let $k \in \mathbb{N}$. Note that by Proposition 4, the probability that $N(P_k(I)) \leq L^{1/3}$ where $I$ is a term of $D$ is,

$$\frac{|S(x)|}{L^{1.5+o(1)}|S(x)|} = \frac{1}{L^{1.5+o(1)}}.$$

As $x \to \infty$, this probability tends to 0. Therefore, for any $k \in \mathbb{N}$ and each term $I$ of $D$, we may assume that $N(P_k(I)) > L^{1/3}$.

For ease of explanation, let $Pr([L^{\sqrt{2}-\epsilon}])$ denote the probability that a sequence of length $[L^{\sqrt{2}-\epsilon}]$ has a square dependent subsequence $D$. Then what we have just shown is that in order for $Pr([L^{\sqrt{2}-\epsilon}])$ to *not* tend to 0 as $x \to \infty$, it must be the case that each of the ideals in the square dependent subsequence $D$ are $L^3$-smooth, and the norm of any prime ideal divisor of an ideal of $D$ must be greater than $L^{1/3}$. We will now give an additional necessary condition for a sequence of length $[L^{\sqrt{2}-\epsilon}]$ to have a square dependent subsequence $D$.

Let $k \in \mathbb{N}$. Choose $I_0 \in D$ such that $N(P_k(I_0))$ is maximal. By Proposition 6, the probability that $P_i(I_0)^2 \mid I_0$ for $1 \leq i \leq k$ is at most

$$\frac{|S(x)|}{L^{\sqrt{2}+o(1)}|S(x)|} = \frac{1}{L^{\sqrt{2}+o(1)}},$$

which tends to 0 as $x \to \infty$. Therefore, we may assume that $P_1(I_0), P_2(I_0), \ldots P_k(I_0)$ are distinct prime ideals of $\mathcal{O}_K$.

Since $I_0$ is involved in the square dependency, we know that there are ideals $I_1, I_2, \dots, I_k$ of our sequence of length $[L^{\sqrt{2}-\epsilon}]$, where we take $I_j = \mathcal{O}_K$ for $j > [L^{\sqrt{2}+\epsilon}]$, such that $P_1(I_0)P_2(I_0)\cdots P_k(I_0) \mid I_1 I_2 \cdots I_k$. Note that it may be the case that for some $j \in \{1, 2, \dots, k\}$ we have $P_i(I_0) \nmid I_j$ for all $i \in \{1, 2, \dots, k\}$, that is, we may not need $k$ terms of our original sequence to obtain a multiple of $P_1(I_0)P_2(I_0)\cdots P_k(I_0)$. However, we know that we need at most $k$ terms from the original sequence by the simple argument from linear algebra given in chapter 1.

So given a that our sequence of length $[L^{\sqrt{2}-\epsilon}]$ has a square dependent subsequence, for any $k \in \mathbb{N}$ we can determine a set of $k$ prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_k$ with $L^{1/3} < N(\mathfrak{p}_k) \leq \cdots \leq N(\mathfrak{p}_1) \leq L^3$ and $N(\mathfrak{p}_1 \cdots \mathfrak{p}_k) \leq x$. From these prime ideals, we deduced that our sequence must contain a $(k+1)$-tuple $I_0, I_1, I_2, \dots I_k$, such that $\mathfrak{p}_1 \cdots \mathfrak{p}_k \mid I_0$ and $\mathfrak{p}_j \mid I_j$ for $1 \leq j \leq k$. We will show the probability that $\mathcal{D}$ contains such a $(k+1)$-tuple tends to 0 as $x \to \infty$. This will imply the probability that $\mathcal{D}$ has a square dependent subsequence also tends to 0 as $x \to \infty$. We begin by determining an upper bound on the number of $(k+1)$-tuples of the form described above.

For a given $k \in \mathbb{N}$, let $\mathcal{P}_k$ denote the set of all possible $k$-tuples of distinct prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_k$, such that $L^{1/3} < N(\mathfrak{p}_k) \leq \cdots \leq N(\mathfrak{p}_1) \leq L^3$ and $N(\mathfrak{p}_1 \cdots \mathfrak{p}_k) \leq x$. For a given $k$-tuple $\mathfrak{p}_1, \dots, \mathfrak{p}_k$ of $\mathcal{P}_k$, let $\mathcal{F}(\mathfrak{p}_1, \dots, \mathfrak{p}_k)$ denote the set of $(k+1)$-tuples $M_0, M_1, \dots, M_k \subseteq \mathcal{O}_k$ where $M_0 = \mathfrak{p}_1 \cdots \mathfrak{p}_k$ and $M_1 \cdots M_k$ is an ordered factorization of $\mathfrak{p}_1 \cdots \mathfrak{p}_k$, that is, $\mathfrak{p}_1 \cdots \mathfrak{p}_k = M_1 \cdots M_k$, where some $M_i$ may be the entire ring $\mathcal{O}_k$. For a given $(k+1)$-tuple $(M_0, M_1, \dots, M_k) \in \mathcal{F}(\mathfrak{p}_1, \dots, \mathfrak{p}_k)$, let $\mathcal{C}(M_0, M_1, \dots, M_k)$ denote the set of $(k+1)$-tuples $I_0, I_1, \dots, I_k \subseteq \mathcal{O}_K$ such that $I_i$ is $L^3$-smooth, $L^{1/3} < N(P_k(I_i)) \leq N(\mathfrak{p}_k)$, and $M_i \mid I_i$ for all $i$. Note that $(k+1)$-tuples of this form are exactly those whose existence is a necessary condition for our sequence $\mathcal{D}$ to have a square dependent subsequence.

Hence, the total number of $(k+1)$-tuples $I_0, I_1, \dots, I_k \subseteq \mathcal{O}_K$ such that $I_i$ is

$L^3$-smooth, $L^{1/3} < N(P_k(I_i)) \leq N(\mathcal{P}_k)$, and $M_i \mid I_i$ for $1 \leq i \leq k$ is

$$\sum_{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k} \; \sum_{(M_0,M_1,\ldots,M_k)\in\mathcal{F}(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)} |\,\mathcal{C}(M_0, M_1, \ldots, M_k)\,|. \tag{3.22}$$

Let $T(i) = (L^{a_{i-1}}, L^{a_i}]$ for $i = 1, 2, \ldots, \lceil (8/3)\ln(L) \rceil$, as in the proof of Proposition 6. Then we can write (3.22) as

$$\sum_{i=1}^{\lceil (8/3)\ln(L)\rceil} \; \sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k \\ N(\mathfrak{p}_k)\in T(i)}} \; \sum_{(M_0,M_1,\ldots,M_k)\in\mathcal{F}(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)} |\,\mathcal{C}(M_0, M_1, \ldots, M_k)\,|.$$

Let $1 \leq i \leq \lceil (8/3)\ln(L) \rceil$. For any $(k+1)$-tuple $(M_0, M_1, \ldots, M_k) \in \mathcal{F}(\mathfrak{p}_1, \ldots, \mathfrak{p}_k)$ we know that $N(M_j) \leq N(\mathfrak{p}_1 \cdots \mathfrak{p}_k) \leq L^{3k}$ for $0 \leq j \leq k$. Hence, by Proposition 5, for $0 \leq j \leq k$, the number of $L^3$-smooth ideals $I$ with norm at most $x$, having the properties that $M_j \mid I$ and $N(P_k(I)) \leq L^{a_i}$ is at most $|S(x)|/(N(M_j)L^{1/(2a_i)+o(1)})$ for $0 \leq j \leq k$. Hence,

$$\sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k \\ N(\mathfrak{p}_k)\in T(i)}} \; \sum_{(M_0,M_1,\ldots,M_k)\in\mathcal{F}(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)} |\,\mathcal{C}(M_0, M_1, \ldots, M_k)\,|.$$

$$\leq \sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k \\ N(\mathfrak{p}_k)\in T(i)}} \; \sum_{(M_0,M_1,\ldots,M_k)\in\mathcal{F}(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)} \left( \frac{|S(x)|^{k+1}}{N(M_0)N(M_1)\cdots N(M_k)L^{(k+1)/(2a_i)+o(1)}} \right)$$

$$= \sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k \\ N(\mathfrak{p}_k)\in T(i)}} \; \sum_{(M_0,M_1,\ldots,M_k)\in\mathcal{F}(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)} \left( \frac{|S(x)|^{k+1}}{N(\mathfrak{p}_1\cdots\mathfrak{p}_k)^2 L^{(k+1)/(2a_i)+o(1)}} \right)$$

As was previously noted, there are $c_k \in \mathbb{N}$ ordered factorizations of $\mathfrak{p}_1 \cdots \mathfrak{p}_k$, where $c_k$ is dependent only on $k$. Writing the constant $c_k$ as $1/L^{o(1)}$ we get,

$$\sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k \\ N(\mathfrak{p}_k)\in T(i)}} \; \sum_{(M_0,M_1,\ldots,M_k)\in\mathcal{F}(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)} \left( \frac{|S(x)|^{k+1}}{N(\mathfrak{p}_1\cdots\mathfrak{p}_k)^2 L^{(k+1)/(2a_i)+o(1)}} \right)$$

$$= \sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k \\ N(\mathfrak{p}_k)\in T(i)}} \left(\frac{1}{L^{o(1)}}\right)\left(\frac{|S(x)|^{k+1}}{N(\mathfrak{p}_1\cdots\mathfrak{p}_k)^2 L^{(k+1)/(2a_i)+o(1)}}\right)$$

$$= \sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k \\ N(\mathfrak{p}_k)\in T(i)}} \left(\frac{|S(x)|^{k+1}}{N(\mathfrak{p}_1\cdots\mathfrak{p}_k)^2 L^{(k+1)/(2a_i)+o(1)}}\right)$$

$$= \left(\frac{|S(x)|^{k+1}}{L^{(k+1)/(2a_i)+o(1)}}\right) \sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k \\ N(\mathfrak{p}_k)\in T(i)}} \frac{1}{N(\mathfrak{p}_1\cdots\mathfrak{p}_k)^2}$$

Let us consider just the sum for a moment. Note that,

$$\sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k \\ N(\mathfrak{p}_k)\in T(i)}} \frac{1}{N(\mathfrak{p}_1\cdots\mathfrak{p}_k)^2} \leq \left(\sum_{N(\mathfrak{p}_k)\in T(i)} \frac{1}{N(\mathfrak{p}_k)^2}\right)\left(\sum_{L^{a_{i-1}}\leq N(\mathfrak{p})<L^3} \frac{1}{N(\mathfrak{p})^2}\right)^{k-1}.$$

By (3.21) in the proof of Proposition 6 we have,

$$\sum_{N(\mathfrak{p}_k)\in T(i)} \frac{1}{N(\mathfrak{p}_k)^2} = \frac{1}{L^{a_i+o(1)}}.$$

Hence,

$$\left(\sum_{N(\mathfrak{p}_k)\in T(i)} \frac{1}{N(\mathfrak{p}_k)^2}\right)\left(\sum_{L^{a_{i-1}}\leq \mathfrak{p}<L^3} \frac{1}{N(\mathfrak{p})^2}\right)^{k-1}$$

$$= \left(\frac{1}{L^{a_i+o(1)}}\right)\left(\sum_{L^{a_{i-1}}\leq \mathfrak{p}<L^3} \frac{1}{N(\mathfrak{p})^2}\right)^{k-1}.$$

Next we note that,

$$\left(\sum_{L^{a_{i-1}}\leq \mathfrak{p}<L^3} \frac{1}{N(\mathfrak{p})^2}\right)^{k-1} \leq \left(\frac{n}{L^{a_{i-1}}} \sum_{L^{a_{i-1}}\leq p<L^3} \frac{1}{p}\right)^{k-1}$$

$$= \left(\frac{n}{L^{a_{i-1}}}\right)^{k-1} \left(\ln\ln(L^3) - \ln\ln(L^{a_{i-1}}) + o(1)\right)^{k-1}$$

$$= \frac{n^{k-1}\left(\ln(3) - \ln(a_{i-1}) + o(1)\right)^{k-1}}{L^{(k-1)a_{i-1}}},$$

with the second equality following from the fact that $\sum_{p\leq y} = \ln\ln(y) + B + o(1)$, with $B \approx 0.2615$ [3, Thm. 1.4.2]. Since,

$$(k-1)a_{i-1} \;=\; (k-1)\left(\frac{1}{3}+\frac{(i-1)}{\ln(L)}\right) = (k-1)a_i - \frac{(k-1)}{\ln(L)}.$$

we have,

$$\frac{n^{k-1}\left(\ln(3)-\ln(a_{i-1})+o(1)\right)^{k-1}}{L^{(k-1)a_{i-1}}}$$

$$= \;\left(\frac{1}{L^{(k-1)a_i}}\right) e^{k-1} n^{k-1}\left(\ln(3)-\ln(a_{i-1})+o(1)\right)^{k-1}.$$

Since $a_{i-1} = 1/3 + (i-1)/\ln(L) \to 1/3$ as $x \to \infty$, we have

$$e^{k-1} n^{k-1}\left(\ln(3)-\ln(a_{i-1})+o(1)\right)^{k-1} \;=\; \frac{1}{L^{o(1)}}.$$

Therefore,

$$\left(\frac{1}{L^{(k-1)a_i}}\right) e^{k-1} n^{k-1}\left(\ln(3)-\ln(a_{i-1})+o(1)\right)^{k-1} \;=\; \frac{1}{L^{(k-1)a_i+o(1)}}.$$

Putting this all together, we have

$$\sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k\\ N(\mathfrak{p}_k)\in T(i)}} \frac{1}{N(\mathfrak{p}_1\cdots\mathfrak{p}_k)^2} \;\leq\; \left(\sum_{N(\mathfrak{p}_k)\in T(i)} \frac{1}{N(\mathfrak{p}_k)^2}\right)\left(\sum_{L^{a_{i-1}}\leq\mathfrak{p}<L^3} \frac{1}{N(\mathfrak{p})^2}\right)^{k-1}$$

$$\leq\; \left(\frac{1}{L^{(a_i+o(1)}}\right)\left(\frac{1}{L^{(k-1)a_i+o(1)}}\right) \;=\; \frac{1}{L^{ka_i+o(1)}}$$

which implies,

$$\left(\frac{|S(x)|^{k+1}}{L^{(k+1)/(2a_i)+o(1)}}\right) \sum_{\substack{(\mathfrak{p}_1,\ldots,\mathfrak{p}_k)\in\mathcal{P}_k\\ N(\mathfrak{p}_k)\in T(i)}} \frac{1}{N(\mathfrak{p}_1\cdots\mathfrak{p}_k)^2} \;<\; \frac{|S(x)|^{k+1}}{L^{ka_i+(k+1)/(2a_i)+o(1)}}.$$

Next we will determine an upper bound for $|S(x)|^{k+1}/L^{ka_i+(k+1)/(2a_i)+o(1)}$ which holds for all $1 \leq i \leq \lceil (8/3)\ln(L) \rceil$. To do so we note that for any $1 \leq i \leq \lceil (8/3)\ln(L) \rceil$, the expression $|S(x)|^{k+1}/L^{ka_i+(k+1)/(2a_i)+o(1)}$ is maximized when $ka_i + (k+1)/2a_i$ is minimized. This occurs when $a_i = \sqrt{(k+1)/2k}$, giving a minimum value of $\sqrt{2k(k+1)}$. Hence,

$$\frac{|S(x)|^{k+1}}{L^{ka_i+(k+1)/(2a_i)+o(1)}} \;\leq\; \frac{|S(x)|^{k+1}}{L^{\sqrt{2k(k+1)}+o(1)}}. \tag{3.23}$$

for all $1 \leq i \leq \lceil (8/3)\ln(L) \rceil$.

All that remains is to sum (3.23) for all $i$. As was shown in the proof of Proposition 6, $\lceil (8/3)\ln(L) \rceil = 1/L^{o(1)}$. Hence,

$$\sum_{i=1}^{\lceil (8/3)\ln(L) \rceil} \frac{|S(x)|^{k+1}}{L^{\sqrt{2k(k+1)}+o(1)}} \;=\; \left(\frac{1}{L^{o(1)}}\right) \frac{|S(x)|^{k+1}}{L^{\sqrt{2k(k+1)}+o(1)}} \;=\; \frac{|S(x)|^{k+1}}{L^{\sqrt{2k(k+1)}+o(1)}}.$$

The above gives us an upper bound on the number of possible $(k+1)$-tuples whose existence is required for our sequence of length $[L^{\sqrt{2}-\epsilon}]$ to have a square dependent subsequence. Therefore, by Proposition 1, the probability that our sequence of length $L^{\sqrt{2}-\epsilon}$ contains at least one of the desired $(k+1)$-tuples is at most

$$\frac{L^{(k+1)(\sqrt{2}-\epsilon)}|S(x)|^{k+1}}{|S(x)|^{k+1}L^{\sqrt{2k(k+1)}+o(1)}} \;=\; L^{(k+1)(\sqrt{2}-\epsilon)-\sqrt{2k(k+1)}+o(1)}.$$

Thus far, our choice of $k \in \mathbb{N}$ has been arbitrary, however, we should note that $k$ can be at most $[L^{\sqrt{2}-\epsilon}]$, as the probability of having a $([L^{\sqrt{2}-\epsilon}]+1)$-tuple in a sequence of length $[L^{\sqrt{2}-\epsilon}]$ is obviously 0. Choose $k \in \mathbb{N}$ such $\sqrt{2k/(k+1)} > \sqrt{2} - \epsilon$. To see that such a choice is justified, note that

$$\lim_{k\to\infty} \sqrt{\frac{2k}{k+1}} \;=\; \sqrt{2}.$$

In choosing $k$ in such a manner, we see that

$$(k+1)(\sqrt{2}-\epsilon) - \sqrt{2k(k+1)} \;=\; (k+1)(\sqrt{2}-\epsilon-\sqrt{2k/(k+1)}) < 0$$

Hence, the probability that our sequence of length $[L^{\sqrt{2}-\epsilon}]$ contains at least one of the desired $(k+1)$-tuples, which is given by

$$= \; L^{(k+1)(\sqrt{2}-\epsilon)-\sqrt{2k(k+1)}+o(1)},$$

tends to 0. As noted previously, this implies the probability of our sequence of length $[L^{\sqrt{2}-\epsilon}]$ having a square dependent subsequence tends to 0 as well.

■

# References

[1] J.A. Buchmann and C.S. Hollinger. On smooth ideals in number fields. *Journal of Number Theory*, 59:82–87, 1996.

[2] M.R. Canfield, P. Erdös, and C. Pomerance. On a problem of Oppenheim concerning "factorisatio numerorum". *Journal of Number Theory*, 17:1–28, 1983.

[3] R. Crandall and C. Pomerance. *Prime Numbers: A Computational Perspective.* Springer, 2005.

[4] E. Croot, A. Granville, R. Premantle, and P. Tetali. Sharp transitions in making squares. http://www.math.gatech.edu/∼ecroot/, 2008.

[5] N.G. de Bruijn. The asymptotic behaviour of a function occurring in the theory of primes. *J. of Indian Math. Soc.*, pages 25–32, 1951.

[6] A. Granville. Smooth numbers: Computational number theory and beyond. *Algorithmic number theory*, 44:1–57, 2008.

[7] A. Hildebrand and G. Tenenbaum. Integers without large prime factors. *Journal de Théorie des Nombres de Bordeaux*, 5(2):411–484, 1993.

[8] P. Moree and C.L. Stewart. Some Ramanujan-Nagell equations with many solutions. *Nederl. Akad. Wetensch. Proc.*, pages 465–472, 1990.

[9] M. Ram Murty and J. Van Order. Counting integral ideals in a number field. *Expositiones Mathematicae*, 25(1):53–66, 2007.

[10] N.G.de Bruijn. On the number of positive integers $\leq x$ and free of prime divisors $> y$, II. *Nederl. Akad. Wetensch. Proc.*, A,69:240–247, 1966.

[11] C. Pomerance. The number field sieve. *Mathematics of Computation, 1943-1993, Fifty Years of Computational Mathematics : Proceedings of Symposia in Applied Mathematics*, 48:465–480, 1994.

[12] C. Pomerance. The role of smooth numbers in number theoretic algorithms. *Proceeding of the International Congress of Mathematicians, Zürich, Switzerland,* 1:411–422, 1994.

[13] C. Pomerance. Multiplicative independence for random integers. *Analytic Number Theory, Proceedings of a Conference in Honor of Heini Halberstam*, pages 703–710, 1996.

[14] V. Ramaswami. On the number of positive integers less than $x$ and free of prime divisors greater that $x^c$. *Bulletin of the American Mathematical Society*, 55:1122–1127, 1949.

[15] E.J. Scourfield. On ideals free of large prime factors. *Journal de Théorie des Nombres de Bordeaux*, 16(3):733–772, 2004.

[16] W. Sierpiński. *Prace Mat. Fiz.*, pages 77–118, 1906.