ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

# Average Case Error Estimates of the Strong Lucas Probable Prime Test

Master Thesis

Semira Einsele

Monday 27th July, 2020

Advisors: Prof. Dr. Kenneth Paterson
Prof. Dr. Özlem Imamoglu

Department of Mathematics, ETH Zürich

**Abstract**

Generating and testing large prime numbers is crucial in many public-key cryptography algorithms. A common choice of a probabilistic primality test is the strong Lucas probable prime test which is based on the Lucas sequences. Throughout this work, we estimate bounds for average error behaviour of this test.

To do so, let us consider a procedure that draws $k$-bit odd integers independently from the uniform distribution, subjects each number to $t$ independent iterations of the strong Lucas probable prime test with randomly chosen bases, and outputs the first number that passes all $t$ tests. Let $q_{k,t}$ denote the probability that this procedure returns a composite number. We show that $q_{k,1} < \log(k)k^2 4^{2.3-\sqrt{k}}$ for $k \geq 2$. We see that slightly modifying the procedure, using trial division by the first $l$ odd primes, gives remarkable improvements in this error analysis. Let $q_{k,l,t}$ denote the probability that the now modified procedure returns a composite number. We show that $q_{k,128,1} < k^2 4^{1.87727-\sqrt{k}}$ for $k \geq 2$. We also give general bounds for both $q_{k,t}$ and $q_{k,l,t}$ when $t \geq 2, k \geq 21$ and $l \in \mathbb{N}$.

In addition, we treat the numbers, that add the most to our probability estimate differently in the analysis, and give improved bounds for large $t$. With the goal of doing a similar analysis, we characterize the numbers with the second largest contribution to the probability estimate.

Moreover, it is known that every odd composite integer $n$ that is not a product of twin primes will be declared prime at most $4n/15$ times. Although with this result we cannot directly conclude that $q_{k,t} \leq (4/15)^t$, we indeed show that $q_{k,t} \leq (4/15)^t$ for $k \geq 111$ .

**Acknowledgements**

First and foremost, I would like to thank my direct supervisor, Prof. Kenny Paterson, for his excellent supervision and for all the work he put into this project. I also want to thank him for helping me to choose a topic, which turned out to be fascinating, for always being available for questions and discussions, regularly providing constructive feedback and for continuing to be a constant source of inspiration throughout the project.

Second, I would like to thank Prof. Özlem Imamoglu for enabling and supporting this collaboration by agreeing to supervise my work.

Last but not least, I would like to thank Mia Filić for all the fruitful discussions, her countless exciting inputs and her always friendly and supportive attitude.

# Contents

Chapter 1

# Introduction

Most modern cryptosystems make use of large prime numbers, either as secret or public parameters. For example, the public-key cryptosystem RSA, which is commonly used in many commerce web sites, is based on the simple number theoretic fact that it should be easy to find and multiply large prime numbers, but it should be very difficult to factor a large composite into primes. Prime generation is therefore a basic cryptographic operation. A way to generate large prime numbers is to choose integers of appropriate size at random and then test them for primality, until a prime number is found. This encourages us to search for primality testing algorithms that are polynomial in complexity. There exist several sophisticated general-purpose algorithms that deterministically test primality, but their efficiency is not sufficient for most applications. In practice, one therefore resorts to *probable prime* tests, and therefore allows a small probability of letting a composite number pass as a prime. In this work, we will simply refer to both probabilistic and deterministic tests as primality tests.

For a variety of applications, such as prime generation, it is important to know how the test behaves in the average case, that is, which error probability do we expect when it is known that the input has been chosen according to some particular probability distribution. There are also other scenarios where the public key parameters, such as the Diffe-Hellman key exchange protocol, must be verified. Since these parameters may have been chosen by an adversary, it is important that the worst-case error probability of the primality test is small.

In the seventeenth century, Fermat came up with a theorem referred to as Fermat's little theorem, which is the basis of the Fermat primality test. However, because there exist numbers, that always fail the Fermat primality test, so called Carmichael numbers, other more powerful extensions of the test have been studied. In 1976 Miller [12] described a true primality test that determines whether a given number is prime in runtime bounded

by $polylog(n)$, where $n$ is the number tested for primality. Miller's test is conditional, meaning that it relies on an unproven hypothesis; in this case the extended Riemann hypothesis. This is widely believed to be true, but mathematicians have been trying to prove it for over one hundred years, yet unsuccessful. In the same year, Rabin [19], [20] modified Miller's algorithm to present it as an unconditional, but randomized algorithm. Rabin's algorithm is always correct when inputting primes, and has a non-zero error probability when the input is composite. Fortunately, this probability of error can be made arbitrarily small. The modified algorithm is commonly referred to as the Miller-Rabin primality test and is the most commonly used primality test in practice today. Shortly after the discovery of Miller, Solovay and Strassen [23] discovered an alternative probabilistic algorithm for testing primality with properties similar to the Miller-Rabin test, which is also employed today. The test resembles Fermat's test, but it does not have the drawback of having composites which are always declared prime. Furthermore, the result is unconditional. In 1980 Baillie and Wagstaff [18] introduced another probabilistic primality test, the Lucas test and its stricter variant the strong Lucas test, which is based on the Lucas sequences. In 1999, Agrawal and Biswas [2] gave a new type of randomized primality test, which exploits a different number theoretic generalization of Fermat's Little Theorem. In 2002 Agrawal, Kayal, and Saxena [11] described the first known unconditional algorithm, the AKS primality test, which can provably determine primality, but is polynomial in complexity. Their algorithm has been a major breakthrough as they essentially showed that primality testing belongs to the complexitiy class $\mathcal{P}$. It proceeds by derandomizing the algorithm proposed in 1999. However, tests like the AKS test are only of theoretical interest because they are too inefficient to be useful in practice. In contrast, tests that accept composite numbers with bounded probability are typically implemented much more frequently.

When studying primality tests, one easily sees that nearly all known primality tests are built on the same basic principle: from the input number $n$, one defines an Abelian group and then tests whether the group structure we expect to see in case $n$ is prime is actually present. The Fermat, Miller-Rabin, Solovay-Strassen and AKS primality tests all use the group $(\mathbb{Z}/n\mathbb{Z})^\times$ in exactly this way. A natural alternative is to try a quadratic extension of $\mathbb{Z}_n$, that is, we look at the ring $\mathbb{Z}_n[x]/(f(x))$ where $f(x)$ is a degree 2 polynomial chosen such that it is guaranteed to be irreducible if $n$ is prime. In that case, the ring is isomorphic to the finite field with $n^2$ elements, $GF(n^2)$. Both the Lucas test as well as the strong Lucas test are based on this approach.

The introduction of Lucas sequences in primality testing opens up more directions in primality testing research. In Chapter 2, we introduce the main concepts, with an emphasis on the Lucas sequences. In Chapter 3, we look at the analog of the Rabin-Monier theorem for the strong Lucas test, which

quantifies the worst-case error probability. The worst-case numbers is the set of odd composite numbers that have the highest probability of being classified as prime. Then, we proceed to analyze the worst-case numbers of the strong Lucas test. As already mentioned above, the average case error analysis is of major importance. For this, consider an algorithm that repeatedly chooses random $k$-bit number and subjects each number to $t$ iterations of the strong Lucas test. If the chosen number passes all $t$ tests, then the procedure will return that number; otherwise another $k$-bit integer is selected and then tested. The procedure ends when a number that passes all $t$ tests is found. Let $q_{k,t}$ denote the probability that such a number is composite. In Chapter 4, we establish numeric upper bounds for $q_{k,t}$. We realize that by modifying the procedure slightly, using trial division by small primes prior to running the strong Lucas test, we get notable improvements of the bounds. Using trial division is a common assumption as it is quite frequently used prior to more expensive tests in cryptographic software. We then treat the numbers that add the most to our probability estimate differently, enabling us to establish bounds that are good when $t$ is large. In Chapter 5 we classify the numbers that contribute the next most to our estimate. It is known that every odd composite integer $n$ that is not a product of twin primes will be declared prime at most $4n/15$ times. Even though with this result we cannot directly conclude that $q_{k,t} \leq (4/15)^t$, we will indeed show in Chapter 6, we have $q_{k,t} \leq (4/15)^t$ for $k \geq 111$.

Chapter 2

# Preliminaries

## 2.1 Primality Tests, Probable Prime Tests and Pseudoprimes

Suppose $S$ is an easily checkable arithmetic statement and we have a theorem "If $n$ is prime, then $S$ is true about $n$". If we are presented with a large number $n$, and we wish to decide whether $n$ is prime or composite, we may try out the arithmetic statement $S$ and see whether it actually holds for $n$. If the statement is not true, we have proved that $n$ is composite. If the statement holds, however, it may be that $n$ is prime, and it also may be that $n$ is composite. Therefore, we have the notion of an *S-pseudoprime*, which is a composite integer for which $S$ holds. Since the test using $S$ does not deterministically show primality of $n$, we cannot call it a *primality test*, but a *probable prime test*, as it can falsely identify a composite number as prime. However, usually the error probability of such a test is extremely small, whereas the running time is a lot faster than the deterministic ones, making it very applicable in practice. Therefore, such probabilistic primality tests are often just called primality test.

One example might be the theorem, If $n > 2$ is prime, then $n$ is odd. This arithmetic property is easily checked for any given input $n$. However, as one can see, this test is not very strong evidence of primality, as far more pseudoprimes exist for this test than genuine primes. Thus, for the concept of "pseudoprime" to be useful, it will have to be the case that there are, in some appropriate sense, few of them.

## 2.2 The Fermat Primality Test

The fact that the residue $a^b \bmod n$ may be rapidly computed due to modular exponentiation is fundamental to many algorithms in number theory. Not

least of these is the exploitation of Fermat's Little Theorem as a means to distinguish between primes and composites. Due to its simple nature, Fermat's Little Theorem is among one of the most studied primality tests.

**Theorem 2.1 (Fermat's Little Theorem)** *Let $p$ be a prime number. For all $a$ relatively prime to $n$, we have*

$$a^{p-1} \equiv 1 \bmod p.$$

The theorem is not true for composite numbers in general, hence it is a good tool to investigate primality. Basically, to test whether $p$ is prime, we can check if a randomly chosen integer $a$ coprime to $p$ satisfies Fermat's Little Theorem. This procedure is called the Fermat test. If the theorem does not hold for $a$ and $p$, we can be sure that $p$ is composite, and thus the test is completed. However, if $a$ and $p$ do satisfy the theorem, we cannot necessarily be convinced that $p$ is prime, as the congruence can also be true for integers that are composite. This leads us to the following definition.

**Definition 2.2** *A* pseudoprime base *$a$ or* $\mathrm{psp}(a)$ *is a composite number $n$, such that $a^{n-1} \equiv 1 \bmod n$, i.e. it satisfies Fermat's Little Theorem using base $a$, even though $n$ is not prime.*

One could speculate if it was enough to verify this for all $a$ which are relatively prime to $n$. Unfortunately, this is not the case. There are many composite numbers, called Carmichael numbers, that pass the Fermat test for every base $a$ coprime to $n$. The smallest one is 561.

**Definition 2.3** *A composite integer $n$ is called* Carmichael number *if it satisfies the congruence $a^{n-1} \equiv 1 \bmod n$ for all integers $a$, which are relatively prime to $n$.*

Ahlford, Granville, and Pomerance [25] proved in 1994 that there are infinitely many Carmichael numbers. An alternative and equivalent definition of Carmichael numbers is given by Korselt's criterion.

**Theorem 2.4 (Korselt's Criterion)** *A positive composite integer $n$ is a Carmichael number if and only if $n$ is square-free, and for all prime divisors $p$ of $n$, it is true that $p-1 \mid n-1$.*

Therefore, we are interested in the following question: Given $n$, how many bases $a$ satisfy Fermat's Little Theorem? Let us denote the set

$$\mathscr{F}(n) = \{a \bmod n \mid n \text{ is a pseudoprime base } a\}$$

and its cardinality $F(n) = \mid \mathscr{F}(n) \mid$ . There is a precise formula for $F(n)$, the proof can be found in [18].

**Theorem 2.5** *Let $n = \prod_{i=1}^{s} p_i^{r_i}$ be a positive integer. The number of bases $a \bmod n$, for which $n$ is a $psp(a)$ is*

$$F(n) = \prod_{i=1}^{s} (n-1, p_i - 1).$$

Since only those $a$ relatively prime to $n$ are candidates and the Euler's $\varphi$ function counts exactly the numbers $0 < a < n$ which are relatively prime to $n$, this product can be at most $\varphi(n)$. We also see that this product is largest whenever $p_i - 1 \mid n - 1$ for all $i$. If $n$ is square-free, then $\prod_{i=1}^{s}(n-1, p_i - 1) = \varphi(n)$. Thus $n$ must be a Carmichael number.

Carmichael numbers are rare, and if a number $x$ is chosen at random, it is unlikely to be one. For each positive integer $x$, let $C(x)$ be the number of Carmichael numbers that are less than $x$. Pomerance showed in [16] and [15] that $\exp(\log(x)^{\frac{15}{37}}) \leq C(x) \leq x \cdot \exp\left(\frac{\log(x)\log(\log(\log(x)))}{\log(\log(x))}\right)$. However, they are not rare enough to be ignored completely.

Even though the Fermat primality test is not used in practice, let us consider its error bound. Let $n \leq x$ be an odd integer which is chosen uniformly at random, let $X$ denote the event that $n$ is composite and let $Y$ be the event that $n$ is a probable prime base $a$, where $1 < a < n$ is also chosen uniformly at random. It has been shown in [22] that for $x \geq 10^{10^5}$, we have $\mathbb{P}[X \mid Y] \leq (\log(x))^{-197}$.

## 2.3 The Miller-Rabin Primality test

Because of the existence of Carmichael numbers which can never be detected as composites by the Fermat based test, slight modifications can eliminate the possibility of Carmichael numbers. One of the most widely used primality tests is the Miller-Rabin test, which is based on the following theorem:

**Theorem 2.6** *Let $n > 1$ be an integer, and write $n - 1 = 2^{\kappa} q$, where $t$ is odd. Then $n$ is a prime if and only if for every $a \not\equiv 0 \bmod n$ one of the following is satisfied*

$$\begin{aligned} &a^q \equiv 1 \bmod n \\ &\textit{or} \\ &\textit{there exists an integer } i < \kappa \textit{ with } a^{2^i q} \equiv -1 \bmod n. \end{aligned} \tag{2.1}$$

If this should hold for some pair $n, a$ we say $n$ is a *strong probable prime base $a$.* Just like pseudoprimes exist for the Fermat test, they also exist for the Miller-Rabin test.

**Definition 2.7 (Strong Pseudoprime)** *We say that $n$ is a strong pseudoprime base $a$, short $spsp(a)$, if $n$ is an odd composite and for the decomposition $n-1 = 2^\kappa q$, with $q$ odd, one of the congruences of (2.1) holds.*

The following is often referred to as "the Miller–Rabin test", which uses Theorem 2.1 with a random choice of $a$.

**Algorithm 1:** Miller-Rabin probabilistic primality test

**Miller-Rabin**$(n)$

**Input:** An odd integer $n > 9$.

**Result:** This probabilistic algorithm attempts to find a witness for $n$ and thus prove that $n$ is composite. If $a$ is a witness, ($n$ is composite) is returned; otherwise, ($n$ is a strong probable prime base $a$) is returned.

Write $n-1 = 2^\kappa q$ with $q$ odd.

Select a random $a \in \{1, \ldots, n-1\}$.

Let $b = a^q \bmod n$.

**if** $b == 1$ *or* $n-1$ **then**
| status = $n$ is strong probable prime base $a$

**else if for** $i \in \{1, \ldots \kappa - 1\}$ **do**
| $b = b^2 \bmod n$

**if** $b == n-1$ **then**

**then**
| status = $n$ is a strong probable prime base $a$

**else**
| status = $n$ is composite

**Return** status

Algorithm 1 is an effective method for recognizing composite numbers and can be used as a way to declare $n$ as prime with sufficiently high probability. Suppose $n$ is a large odd number and we don't know whether $n$ is prime or composite. No one is stopping us from performing the algorithm repeatedly, say 20 times, and fail each time to produce a witness. What should be concluded? Actually, nothing at all can be concluded concerning whether $n$ is prime or composite. The probability that we have failed to produce a witness for a given odd composite is less than $4^{-20}$, which we'll see shortly. This is less than one chance in a trillion. So yes, it is reasonable to strongly conjecture that $n$ is prime. But it has not been proven prime and in fact might not be. However, for practical applications, one may be perfectly happy to

use a number that is almost certainly prime. It is with this mindset that people refer to Algorithm 1 as a "primality test". For more details see [1], Chapter 4.

**Remark 2.8** The Miller-Rabin primality test is used as a probabilistic test. It turns out that if we knew a generalization of the Riemann hypothesis, then we could prove that if $n$ is a strong probable prime base $a$ for the first $2\log(n)$ bases then $n$ is indeed always prime. See [10] for a proof.

Let us introduce the notion of a witness for compositeness.

**Definition 2.9 (Witness for Compositeness)** *If $n$ is a composite number and $a$ is an integer in $\{1,\ldots,n-1\}$, for which (2.1) fails, we say that $a$ is a witness for $n$ using the Miller-Rabin theorem.*

For an odd integer $n$ we denote the set

$$\mathscr{S}(n) = \{a \bmod n \mid n \text{ is a strong pseudoprime base } a\}$$

and its cardinality $S(n) =\mid \mathscr{S}(n) \mid$. The following theorem was provided independently in 1980 by Monier and Rabin, see [14] and [20].

**Theorem 2.10 (The Rabin-Monier Theorem)** *For each odd composite integer $n > 9$ we have*

$$S(n) \leq \frac{1}{4}\varphi(n).$$

For an odd composite integer $n$ at most one quarter of the bases declare $n$ as a strong probable prime. This is a tight bound; there exist odd composite integers $n$ that have exactly $\varphi(n)/4$ such bases. Before we prove Theorem 2.10, we first indicate why it is a significant result. A witness for $n$ is the key to a short proof that $n$ is composite. Theorem 2.10 implies that at least $3/4$ of all integers in $\{1,\ldots,n-1\}$ are witnesses for $n$, when $n$ is an odd composite number. Since one can perform the Miller-Rabin test rapidly, it would seem that it is quite an easy task to produce witnesses for odd composite numbers. Indeed, the probability of Algorithm 1 failing to find a witness in the case of an odd composite number $n$ with $t$ (independent) iterations is less than $(1/4)^t$. So clearly we can make this probability vanishingly small by choosing $t$ large.

The following algorithm may be used for the generation of random numbers

that are likely to be prime.

**Algorithm 2:** Prime Generation

**Input:** The required bitlength $k > 3$ and a security parameter $t \geq 1$.

**Result:** This probabilistic algorithm produces a random $k$-bit (that is, a number in the interval $[2^{k-1}, 2^k)$) strong probable prime; a number that has not been recognized as composite by $t$ iterations of Algorithm 1.

**while** *Candidate not found* **do**
  Choose a random odd integer $n$ in the interval $(2^{k-1}, 2^k)$.

**for** $1 \leq i \leq t$ **do**
  Via Algorithm 1 attempt to find a witness for $n$.; If a witness is found for $n$, candidate not found

**Return** $n$

In order to prove Theorem 2.10, we need some results. In [14] Monier established a formula, which counts for a given $n$ the number of bases $a$ such that $n$ is a $spsp(a)$.

**Theorem 2.11** *Let $p_1^{r_1} \cdot \ldots \cdot p_s^{r_s}$ be the prime decomposition of an odd integer n. We let*

$$\begin{cases} n-1 = 2^\kappa q \\ p_i - 1 = 2^{k_i} q_i \quad \text{for } 0 \leq i \leq s \end{cases} \qquad \text{with } q, q_i \text{ are odd.}$$

*where we have ordered the $p_i$ such that $k_1 \leq \ldots \leq k_s$. The number of bases a in which n is a strong pseudoprime base a is given by the formula*

$$S(n) = (1 + \sum_{j=0}^{k_1-1} 2^{js}) \prod_{i=1}^{s} \gcd(q, q_i). \tag{2.2}$$

With Theorem 2.11, we can easily show the following Lemma:

**Lemma 2.12**

$$\frac{S(n)}{\varphi(n)} \leq \frac{1}{2^{s-1}} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}}.$$

*Moreover if not all the $k_i$ are the same, the following inequality holds:*

$$\frac{S(n)}{\varphi(n)} \leq \frac{1}{2^s} \prod_{i=1}^{s} \frac{\gcd(q, q_i)}{q_i} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}} \leq \frac{1}{2^s} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}}.$$

**Proof** With $\varphi(n) = \prod_{i=1}^{s} p_i^{r_i-1}(p_i - 1) = 2^{k_1+\ldots+k_s} \prod q_i \prod p_i^{r_i-1}$ and Theorem 2.11, we have

$$\frac{S(n)}{\varphi(n)} \leq \frac{1 + \sum_{j=0}^{k_1-1} 2^{js}}{2^{k_1+\ldots+k_s}} \prod_{i=1}^{s} \frac{gcd(q, q_i)}{q_i} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}}.$$

Since $k_1 \leq k_i$ for all $i$, we have that $2^{k_1+\ldots+k_s} \geq 2^{k_1 s}$. Therefore we have

$$\frac{1+\sum_{j=0}^{k_1-1} 2^{js}}{2^{k_1+\ldots+k_s}} \leq \frac{1+\sum_{j=0}^{k_1-1} 2^{js}}{2^{k_1 s}} = \frac{1+\frac{2^{sk_1}-1}{2^s-1}}{2^{sk_1}} = \frac{1-\frac{1}{2^s-1}}{2^{sk_1}} + \frac{1}{2^s-1}.$$

This shows that $\frac{1+\sum_{j=0}^{k_1-1} 2^{js}}{2^{k_1+\ldots+k_s}}$ is a decreasing function in $k_1$. So we can bound it by its value at $k_1 = 1$:

$$\frac{1+\sum_{j=0}^{k_1-1} 2^{js}}{2^{k_1+\ldots+k_s}} \leq \frac{1+\sum_{j=0}^{k_1-1} 2^{js}}{2^{k_1 s}} \leq \frac{2}{2^s} = \frac{1}{2^{s-1}}.$$

Since $\gcd(q, q_i)/q_i \leq 1$ for all $i$, we have proven the first assertion. The second follows in the same way, using $2^{k_1+\ldots+k_s} \geq 2^{sk_1+1}$. □

We are finally ready to prove Theorem 2.10.

**Proof** We distinguish two cases here. Either $n$ is a prime power or it is not. First, let us suppose $n$ is a prime power, meaning $s = 1$. As we have excluded the case $n = 9 = 3^2$, for which the bound $1/3$ holds, either $p_1 \geq 5$ and $r_1 \geq 2$ or $p_1 = 3$ and $r_1 \geq 3$. With Lemma 2.12, which indicates that

$$\frac{S(n)}{\varphi(n)} \leq \frac{1}{p_1^{r_1-1}},$$

the result follows directly.

Now let us suppose that $s = 2$. If $k_1 \leq k_2$, the result immediately follows from the second assertion of Lemma 2.12, which says $\frac{S(n)}{\varphi(n)} \leq \frac{1}{4}$. If $k_1 = k_2$, we have by the first assertion of Lemma (2.12) that

$$\frac{S(n)}{\varphi(n)} = \frac{1}{2}\frac{\gcd(q, q_1)}{q_1}\frac{\gcd(q, 2_2)}{q_2}.$$

At least one of $\gcd(q, q_i)/q_i$ is bounded by $1/3$: If not, since all of $q, q_1, q_2$ are odd, we have that $\gcd(q, q_i) = q_i$ for $i = 1, 2$. This implies that $q_1 \mid q$ and $q_2 \mid q$. Therefore both $q_1$ and $q_2$ divide

$$2^\kappa q = p_1 p_2 - 1 = (2^{k_1} q_1 + 1)(2^{k_2} q_2 + 1) - 1 = 2^{k_1}(q_1 + q_2) + 2^{2k_1} q_1 q_2,$$

this is only possible if $q_1 \mid q_2$ and $q_2 \mid q_1$, which means that $q_1 = q_2$. But this case is already excluded. Therefore

$$\frac{S(n)}{\varphi(n)} \leq \frac{1}{6},$$

which proves the Theorem. □

### 2.3.1 Remarks on estimating $p_{k,t}$

We are interested in the following question: What is the probability that a number produced by Algorithm 2 is composite? We denote this probability by $p_{k,t}$. One might think that Theorem 2.10 immediately answers this question, and that we have $p_{k,t} \leq 4^{-t}$. However, the reasoning is flawed, since it does not take into account the distribution of primes. Let's illustrate this by an example.Suppose $k = 500$ and $t = 1$. We know from the Prime Number Theorem that the probability that a random odd 500-bit number is prime is about 1 chance in 173. Since it is evidently more likely that one will witness an event with probability $1/4$ occurring before an event with probability $1/173$, it may seem that there are much better than even odds that Algorithm 2 will produce composites.

Let $X$ represent the event that $n$ is composite, let $E_i$ denote the event that an integer chosen uniformly at random, say from the set $M_k$ of odd $k$-bit integers, passes the $i$-th round of the Miller-Rabin test and let $Y_t$ denote the event that it passes $t$ consecutive rounds, $Y_t = E_1 \cap E_2 \cap \cdots \cap E_t$, i.e. Algorithm 2 outputs $n$. Theorem 2.10 states that $\mathbb{P}[Y_t \mid X] \leq (\frac{1}{4})^t$. What is relevant, however, to the estimation of $p_{k,t}$ is the quantity $\mathbb{P}[X \mid Y_t]$. Suppose that candidates $n$ are drawn uniformly and randomly from $M_k$. As every prime number passes the test $t$ times we have $\mathbb{P}[Y_t] \geq \mathbb{P}[X^c]$. Then by Bayes' Theorem

$$\mathbb{P}[X \mid Y_t] = \frac{\mathbb{P}[X]\mathbb{P}[Y_t \mid X]}{\mathbb{P}[Y_t]} \leq \frac{\mathbb{P}[Y_t \mid X]}{\mathbb{P}[Y_t]} \leq \frac{1}{\mathbb{P}[Y_t]}\Big(\frac{1}{4}\Big)^t \leq \frac{1}{\mathbb{P}[X^c]}\Big(\frac{1}{4}\Big)^t.$$

Thus the probability $\mathbb{P}[X \mid Y_t]$ may be considerably larger than $(\frac{1}{4})^t$ if $\mathbb{P}[X^c]$ is small, i.e. we assume that the primes in our set of odd integers are scarce. We could construct such an example as follows: for a fixed $t \geq 1$, choose $k$ sufficiently large such that the density of primes in $M_k$ is much less than $4^{-t}$. Assume also that for most composite $m \in M_k$ that the probability that $m$ passes a random bases test is about $1/4$. Then, of course, the probability of it passing $t$ tests is about $4^{-t}$. Suppose that we have an $n$ from $M_k$ that passes $t$ tests. Since we are assuming that the primes in $M_k$ are scarce, it will be much more likely that $n$ is composite rather than prime. So $\mathbb{P}[X \mid Y_t]$ would be close to 1. However, the error-probability of Miller-Rabin is usually far smaller than $(\frac{1}{4})^t$ for all sufficiently large $k$ and it is indeed shown by Burthe [6] that we do have $p_{k,t} \leq 4^{-t}$. He showed that the flawed assumption that led us to the conclusion that the probability of $m$ passing a test $\mathbb{P}[E_i]$ was about $1/4$. In actuality the probability is usually much smaller.

Further refinements for $\mathbb{P}[X \mid Y_t]$ allow some explicit upper bounds on $p_{k,t}$ for various values of $k$ and $t$. If $k$ is large, one gets good results even with $t = 1$ using Algorithm 2. Damgård et al. [8] showed the following results:

**Theorem 2.13** *(i) For $k \geq 2$, we have $p_{k,1} < k^2 4^{2-\sqrt{k}}$.*

*(ii) For* $k \geq 21, 3 \leq t \leq k/9$ *or* $k \geq 88.t = 2$*, we have* $p_{k,t} < k^{3/2}\frac{2^t}{\sqrt{t}}4^{2-\sqrt{tk}}$.

*(iii) For* $k \geq 21$ *and* $t \geq k/9$*, we have* $p_{k,t} < \frac{7}{20}k2^{-5t} + \frac{1}{7}k^{15/4}2^{-k/2-2t} + 12k2^{-k/4-3t}$.

*(iv) For* $k \geq 21$ *and* $t \geq k/4$*, we have* $p_{k,t} < \frac{1}{7}k^{15/4}2^{-k/2-2t}$.

For specific large values of $k$, the paper has even better results, for example $p_{500,1} < 4^{-28}$. Thus, if a randomly chosen odd 500-bit number passes just one iteration of a random Miller-Rabin test, the number is composite with vanishingly small probability, and may be safely accepted as "prime" in all but the most sensitive practical applications.

## 2.4 The Lucas Primality Test

Let $D$, $P$ and $Q$ be integers such that $D = P^2 - 4Q$ is non-zero and $P > 0$. Let $U_0(P,Q) = 0$, $U_1(P,Q) = 1$, $V_0(P,Q) = 2$ and $V_1(P,Q) = P$. The Lucas sequences $U_n(P,Q)$ and $V_n(P,Q)$ associated with the parameters $P$, $Q$ are defined recursively for $n \geq 2$ by

$$U_n(P,Q) = PU_{n-1}(P,Q) - QU_{n-2}(P,Q) \tag{2.3}$$
$$V_n(P,Q) = PV_{n-1}(P,Q) - QV_{n-2}(P,Q). \tag{2.4}$$

Let $\alpha$ and $\beta$ be the distinct roots of the polynomial $X^2 - PX + Q$. We see that $\alpha = \frac{P+\sqrt{D}}{2}$ and $\beta = \frac{P-\sqrt{D}}{2}$. It is also easy to see that

$$\begin{aligned} \alpha\beta &= Q \\ \alpha + \beta &= P \\ \alpha - \beta &= \sqrt{D}. \end{aligned} \tag{2.5}$$

**Lemma 2.14** *(Binet formula)*

$$\begin{aligned} U_n(P,Q) &= \frac{\alpha^n - \beta^n}{\alpha - \beta} \\ V_n(P,Q) &= \alpha^n + \beta^n \quad \forall n \in \mathbb{N}_0. \end{aligned}$$

**Proof** We proceed by induction on $n$.
*Base case* $n = 2$: We have $U_2(P,Q) = PU_1(P,Q) - QU_0(P,Q) = P = \frac{\alpha^2-\beta^2}{\alpha-\beta}$ and $V_2(P,Q) = PV_1(P,Q) - QV_0(P,Q) = P^2 - 2Q = (\alpha+\beta)^2 - 2\alpha\beta = \alpha^2 + \beta^2$.

*Inductive Step*. Let the claim hold for all $k < n$. Then

$$\begin{aligned} U_n(P,Q) &= \frac{\alpha^n - \beta^n}{\alpha - \beta} = \frac{1}{\alpha - \beta}\Big((\alpha+\beta)(\alpha^{n-1} - \beta^{n-1}) - \alpha\beta(\alpha^{n-2} - \beta^{n-2})\Big) \\ &= P(\frac{\alpha^{n-1} - \beta^{n-1}}{\alpha - \beta}) - Q(\frac{\alpha^{n-2} - \beta^{n-2}}{\alpha - \beta}) \\ &= PU_{n-1}(P,Q) - QU_{n-2}(P,Q). \end{aligned}$$

The proof for $V_n(P,Q)$ is identical. □

### 2.4.1 Arithmetic in $\mathbb{Q}[\sqrt{D}]$

We may generalize many of the ideas of the past sections to incorporate finite fields. However, in order to fully grasp all the concepts, we need some theory of quadratic fields.

Let $L, K$ be fields, $K$ a subfield of $L$ and $K/L$ be the field extension. We define the degree of $L/K$ to be the dimension of $L$ as a $K$-vector space and denote it by $[L : K]$.

**Example 2.15** *$\mathbb{Q}[\sqrt{D}]$ has $\mathbb{Q}$-basis $1, \sqrt{D}$. Therefore $[\mathbb{Q}[\sqrt{D}] : \mathbb{Q}] = 2$.*

We define a number field to be a finite extension of $\mathbb{Q}$ and call the degree of a number field $K$ the degree of $[K : \mathbb{Q}]$. We call a number field of degree 2 a quadratic field.

**Lemma 2.16** *Let $K$ be a quadratic field. Then $K = \mathbb{Q}[\sqrt{D}] = \{r + s\sqrt{D} \mid r, s \in \mathbb{Q}\}$, where $D$ is a square-free integer and $D \neq 0, 1$.*

We call $D$ *square-free* if it isn't divisible by any perfect square other than 1, equivalently, $D$ is a product of distinct primes. When working in $\mathbb{Q}[\sqrt{D}]$, it is often useful to assume that $D$ is square-free. This is no loss of generality: if $D' = n^2 D$, then $r + s\sqrt{D'} = r + sn\sqrt{D}$, so $\mathbb{Q}[\sqrt{D'}] = \mathbb{Q}[\sqrt{D}]$.

**Definition 2.17** *An algebraic integer is a complex number which is the root of a monic polynomial $f(x) \in \mathbb{Z}[X]$. We let $\mathcal{O}$ be the set of algebraic integers, namely*

$$\mathcal{O} = \{\alpha \in \mathbb{C} : \exists p(x) \in \mathbb{Z}[X] \text{ monic s.t. } p(\alpha) = 0\}.$$

**Definition 2.18** *If $K$ is a number field, then let*

$$\mathcal{O}_K = \mathcal{O} \cap K = \{\alpha \in K : \exists p(x) \in \mathbb{Z}[X] \text{ monic s.t. } p(\alpha) = 0\}.$$

*We call $\mathcal{O}_K$ the* ring of integers *of $K$. It is a known fact that it is a ring.*

The most famous example is the ring of integers of $\mathbb{Q}$, which is $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$. For this reason we call $\mathbb{Z}$ the set of *rational integers*.

**Definition 2.19** *Let $\alpha \in \mathbb{Q}[\sqrt{D}]$. The* conjugate *of $\alpha = a + b\sqrt{D}$ is $\overline{\alpha} = a - b\sqrt{D}$. We define the* trace *and the* norm *of $\alpha$ by*

$$Tr(\alpha) = \alpha + \overline{\alpha},\ N(\alpha) = \alpha\overline{\alpha}.$$

We are now looking for conditions when an $\alpha \in \mathbb{Q}[\sqrt{D}]$ is in $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$. Let $\alpha \in \mathbb{Q}[\sqrt{D}]$. The conjugate of $\alpha = a + \sqrt{D}b$ is $\bar{\alpha} = a - \sqrt{D}b$. If $\alpha$ is a root of a monic polynomial $p(x) \in \mathbb{Z}[X]$, then we can show that $\overline{\alpha}$ is also a root of $p(x)$: Let $\alpha$ be the root of $p(x) = \sum_{i=0}^{n} c_i x^i$, with $c_i \in \mathbb{Z}$, i.e. $p(\alpha) = \sum_{i=0}^{n} c_i \alpha^i$, then $p(\overline{\alpha}) = \sum_{i=0}^{n} c_i \overline{\alpha}^i = \sum_{i=0}^{n} \overline{c_i \alpha^i} = \overline{\sum_{i=0}^{n} c_i \alpha^i} = \overline{p(x)} = \overline{0} = 0$, which is what we wanted to show. Now $p(x) = q(x)(x - \alpha)(x - \overline{\alpha}) = q(x)(x^2 - (\alpha + \overline{\alpha})x + \alpha\overline{\alpha} = q(x)(x^2 - Tr(\alpha)x + N(\alpha))$, where $Tr(\alpha), N(\alpha) \in \mathbb{Z}$. Thus we see, if $\alpha \in \mathbb{Q}[\sqrt{D}]$ is an algebraic integer, we must have that it is a root of $x^2 - Tr(\alpha)x + N(\alpha))$, where $Tr(\alpha), N(\alpha) \in \mathbb{Z}$. Thus we can rewrite our set of integers in the quadratic field $\mathbb{Q}[\sqrt{D}]$ as:

**Lemma 2.20** *The ring of integers of $\mathbb{Q}[\sqrt{D}]$ is the set*

$$\begin{aligned}\mathcal{O}_{\mathbb{Q}[\sqrt{D}]} &= \{\alpha \in \mathbb{Q} \mid \alpha^2 + b\alpha + c = 0;\ b, c \in \mathbb{Z}\} \\ &= \{\alpha \in \mathbb{Q} \mid Tr(\alpha),\ N(\alpha) \in \mathbb{Z}\}.\end{aligned}$$

The next theorem (see [24]) shows another representation of $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$.

**Theorem 2.21** *Assume $D \in \mathbb{Z}$ is square-free. The ring of integers $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$ is equal to $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]} = \mathbb{Z} + \delta_0\mathbb{Z} = \mathbb{Z}[\delta_0]$, where*

$$\delta_0 = \begin{cases} \sqrt{D} & D \equiv 2, 3 \bmod 4 \\ \frac{1+\sqrt{D}}{2} & D \equiv 1 \bmod 4. \end{cases}$$

**Proof** It is easy to see that $\mathbb{Z} + \delta_0\mathbb{Z} \subseteq \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$, as in either case $\delta_0$ is in $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$: It satisfies the monic equation with coefficients in $\mathbb{Z}$, namely $x^2 - D = 0$ or $x^2 - x + \frac{1-D}{4} = 0$. The latter has coefficients in $\mathbb{Z}$ when $D \equiv 1 \bmod 4$. Now we show the reverse inclusion, $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]} \subseteq \mathbb{Z} + \delta_0\mathbb{Z}$. Let $\alpha = a + \sqrt{D}b \in \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$. By Lemma 2.20 we have that $Tr(\alpha) = 2a, N(\alpha) = a^2 - Db^2 \in \mathbb{Z}$. We put $a = s/2$ for some $s \in \mathbb{N}$, and $b = m/n$ for some $m, n \in \mathbb{Z}$ with $\gcd(m, n) = 1$. Thus we get

$$\begin{aligned}N(\alpha) = (\frac{r}{2})^2 - (\frac{m}{n})^2 D &\iff 4N(\alpha)n^2 = r^2n^2 - 4m^2D \\ &\iff 4m^2D = n^2(r^2 - 4N(\alpha)),\end{aligned}$$

so that $n^2 \mid 4m^2D$. But $\gcd(m, n) = 1$, thus $n^2 \mid m^2D$. If $p$ were an odd prime factor of $n$, we would have $p^2 \mid D$, contradicting the fact that $D$ is square-free.

Thus $n$ has to be a power of 2. Since 2 can be the only even factor of $D$, we must have that $n^2 \mid 8$ or $n^2 \mid 4$. The only $n$ that satisfy this are $n = 1, 2$. In either cases $b = \frac{m}{2}$ for some $m \in \mathbb{Z}$.

With $N(\alpha) = a^2 - Db^2 = \frac{s^2}{4} - \frac{Dm^2}{4} \in \mathbb{Z}$, we have $r^2 = m^2 D \bmod 4$. Since squares can only be 0 or 1 modulo 4, we only have to consider two cases:

1. If $D \not\equiv 1 \bmod 4$, we have that $r^2 \equiv m^2 \equiv 0 \bmod 4$. This implies that $r$ and $m$ are both even integers, hence $a, b \in \mathbb{Z}$, so that $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]} \subseteq \mathbb{Z} + \sqrt{D}\mathbb{Z}$.
2. If $D \equiv 1 \bmod 4$, then $r^2 \equiv m^2 \bmod 4$, which implies $r \equiv m \bmod 2$. Writing $r = m + 2k$ for $k \in \mathbb{Z}$ we see that

$$\alpha = a + b\sqrt{D} = \frac{r + \sqrt{D}m}{2} = \frac{m + 2k + m\sqrt{D}}{2} = k + s\frac{1 + \sqrt{D}}{2}.$$

Thus $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]} \subseteq \mathbb{Z} + \frac{1+\sqrt{D}}{2}\mathbb{Z}$. □

This is another way of saying if $D \not\equiv 1 \bmod 4$, then 1 and $\sqrt{D}$ is an integral basis and if $D \equiv 1 \bmod 4$, then 1 and $\frac{1+\sqrt{D}}{2}$ is an integral basis of $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$.

**Definition 2.22** *The* discriminant *of* $\alpha \in \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$ *is* $disc(\alpha) = (Tr(\alpha))^2 - 4N(\alpha)$.

**Corollary 2.23** *Let* $\alpha, \beta \in \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$. *If* $\mathbb{Z}[\alpha] = \mathbb{Z}[\beta]$, *then* $disc(\alpha) = disc(\beta)$.

This means that $disc(\alpha)$ only depends on the subring $\mathbb{Z}[\alpha] \subseteq \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$ that $\alpha$ generates so it makes sense to write $disc(\mathbb{Z}[\alpha]) = disc(\alpha)$. This leads us to the following definition:

**Definition 2.24** *We put* $D_F = disc(\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}) = disc(\delta_0)$ *and call it the discriminant of the field* $F = \mathbb{Q}[\sqrt{D}]$.

The following table summarizes the basic information about the ring of integers of $F = \mathbb{Q}[\sqrt{D}]$, where $D$ is square-free:

| $D \bmod 4$ | $\delta_0$, where $\mathcal{O}_F = \mathbb{Z}[\delta_0]$ | Eq. for $\delta_0$ | $D_F = disc(\mathcal{O}_F)$ |
|---|---|---|---|
| 2, 3 | $\sqrt{D}$ | $\delta_0^2 - D = 0$ | $4D$ |
| 1 | $\frac{1+\sqrt{D}}{2}$ | ${\delta_0}^2 - \delta_0 + \frac{1-D}{4} = 0$ | $D$ |

**Remark 2.25** We see that we can always write $F = \mathbb{Q}[\sqrt{D_F}]$, as $\sqrt{D_F} = 2\sqrt{D}, \sqrt{D}$, which depends on $D \bmod 4$.

For a rational integer $n$, the ring $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}/(n) = \{x + n\mathbb{Z} : x \in \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}\}$ is a free $\mathbb{Z}/n\mathbb{Z}$-algebra of rank 2. The following results and their proofs can be found in [13].

**Theorem 2.26** *For any quadratic field $\mathbb{Q}[\sqrt{D}]$, $(\mathscr{O}_{\mathbb{Q}[\sqrt{D}]})^\times = \{\alpha \in \mathscr{O}_{\mathbb{Q}[\sqrt{D}]} \mid N(\alpha) = \pm 1\}$, where $(\mathscr{O}_{\mathbb{Q}[\sqrt{D}]})^\times$ is the unit group of $\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}$.*

**Proof** Let $\alpha \in \mathscr{O}_{\mathbb{Q}[\sqrt{D}]}$. If $\alpha$ is a unit, then $\alpha\beta = 1$ for some $\beta \in \mathscr{O}_{\mathbb{Q}[\sqrt{D}]}$. Taking norms on both sides, $N(\alpha)N(\beta) = N(1) = 1$ in $\mathbb{Z}$, so $N(\alpha) = \pm 1$. Conversely, assume $N(\alpha) = \pm 1$. Since $N(\alpha) = \alpha\overline{\alpha}$, we get $\alpha\overline{\alpha} = \pm 1$. Therefore $\pm\overline{\alpha}$, which lies in $\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}$ is an inverse for $\alpha$. □

In the $\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/(n)$-algebra we consider the multiplicative group of norm 1 elements, which we denote by $(\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/(n))^\wedge$. In other words, $(\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/(n))^\wedge$ is the image of the set

$$\{x \in \mathscr{O}_{\mathbb{Q}[\sqrt{D}]} \mid N(x) \equiv 1 \bmod n\}$$

by the canonical map $\mathscr{O}_{\mathbb{Q}[\sqrt{D}]} \to \mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/(n)$.

**Theorem 2.27** *Let $p \nmid 2D$ be a prime number and $r \geq 1$ an integer. The group $(\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/p^r)^\wedge$ is cyclic of order $p^{r-1}(p - (\frac{D}{p}))$.*

We are now ready to connect the parameters $P$ and $Q$ defined through the Lucas sequence and the norm 1 elements $\tau$:

**Proposition 2.28** *Let $D$ be an integer, which is not a perfect square and let $\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}$ be the ring of integers of $\mathbb{Q}[\sqrt{D}]$. Let $n > 1$ be an odd integer, relatively prime to $D$. Then for every integer $P$, there exists an integer $Q$, uniquely determined modulo $n$, such that $P^2 - 4Q \equiv D \bmod n$. Furthermore, the set of integers $P$ such that*

$$\begin{cases} 0 \leq P < n \\ \gcd(P^2 - D, n) = 1 \text{ i.e. } \gcd(Q, n) = 1 \end{cases}$$

*is in one-to-one correspondence with the elements $\tau$ in $(\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/(n))^\wedge$, such that $\tau - 1$ is a unit in $\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/(n)$. I.e.*

$$\{0 \leq P < n : \gcd(P^2 - D, n) =\} \simeq \{\tau \in (\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/n)^\wedge : \tau - 1 \in (\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/n)^\times\}.$$

*This correspondence is expressed by the following formulas*

$$\begin{cases} \tau \equiv (P + \sqrt{D})(P - \sqrt{D})^{-1} \\ P \equiv \sqrt{D}(\tau + 1)(\tau - 1)^{-1} \end{cases} \quad \bmod n\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}.$$

### 2.4.2 Some Lucas Primality tests

In 1980, Baillie and Wagstaff (see [18]) gave a thorough treatment of the use of Lucas sequences in primality testing. They specifically examined the following four congruences:

**Theorem 2.29** *Let $U_p(P,Q)$ and $V_p(P,Q)$ be the Lucas sequences. If $p$ is an odd prime such that $(p, QD) = 1$, then the following congruences all hold:*

$$U_{p-(\frac{D}{p})} \equiv 0 \bmod p, \tag{2.6}$$

$$U_p \equiv \left(\frac{D}{p}\right) \bmod p, \tag{2.7}$$

$$V_p \equiv P \bmod p, \tag{2.8}$$

$$V_{p-(\frac{D}{p})} \equiv 2Q^{(1-(\frac{D}{p}))/2} \bmod p. \tag{2.9}$$

**Proof** Let $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$ be the ring of integers of the field $\mathbb{Q}[\sqrt{D}]$. The quotient ring $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}/(p)$ is isomorphic to either $\mathbb{F}_p \times \mathbb{F}_p$ or $\mathbb{F}_{p^2}$, depending on the Jacobi symbol $(\frac{D}{p})$. We then have the following congruences:

$$\begin{cases} \alpha^p \equiv \alpha \\ \beta^p \equiv \beta \end{cases} \quad \text{when } \left(\frac{D}{p}\right) = 1, \quad \begin{cases} \alpha^p \equiv \beta \\ \beta^p \equiv \alpha \end{cases} \quad \text{when } \left(\frac{D}{p}\right) = -1, \tag{2.10}$$

where the congruences are modulo $p$. In both cases we have

$$\alpha^{p-(\frac{D}{p})} \equiv \beta^{p-(\frac{D}{p})},$$

so that the congruence (2.6) follows directly.
All other congruences (2.9), (2.7) and (2.8) also follow directly with (2.10). □

Sometimes it occurs that an odd positive integer $n$ satisfies one of the congruences (2.6), (2.7), (2.8) or (2.9). This leads us to the following definitions:

**Definition 2.30** *Let $n$ be an odd composite integer. It is called a*

- *Lucas pseudoprimeof of first kind with parameters $P$ and $Q$ if congruence (2.6) holds.*
- *Lucas pseudoprime of second kind with parameters $P$ and $Q$ if congruence (2.7) holds.*
- *Dickson pseudoprime of first kind with parameters $P$ and $Q$ if congruence (2.8) holds.*
- *Dickson pseudoprime of second kind with parameters $P$ and $Q$ if congruence (2.9) holds.*

The congruenecs (2.7) and (2.8) however, are not very useful in primality testing (see [18]): most composite $n$ that satisfy (2.7) have small prime factors; many composite $n$ that satisfy (2.8) are $psp(2)$. Most results about Lucas pseudoprimes refer to congruence (2.6), which seems to be more approachable theoretically. For this reason, they are usually just called *Lucas*

*pseudoprimes with parameters P and Q*, in short *lpsp(P, Q)*. We will summarize some results known about *lpsp(P, Q)*.

For a fixed integer $D$, the number of parameter pairs $(P, Q)$ which lead to a pseudoprime for a given composite $n$ are characterized by the following formula (See [5]), which at first glance seems similar to the formula of Theorem 2.5.

**Theorem 2.31** *Let $D$ be a fixed positive integer and let $n = \prod_{i=1}^{s} p_i^{r_i}$ be a positive odd integer with $\gcd(D, n) = 1$. Then the number of distinct values of $P$ modulo $n$, for which there is a $Q$ such that $P^2 - 4Q \equiv D \bmod n$ and $n$ is a $lpsp(P, Q)$ is*

$$L(D, n) = \prod_{i=1}^{s} \Big[ (n - \left(\frac{D}{n}\right), p_i - \left(\frac{D}{p_i}\right)) - 1 \Big].$$

**Proof** Let $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$ be the ring of integers in $\mathbb{Q}[\sqrt{D}]$. With Proposition 2.28, we know that we can count the elements $\tau$ of $(\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}/(n))^\wedge$ through

$$\tau - 1 \in (\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}/(n))^\times \quad \text{and} \quad \tau^{n-\left(\frac{D}{n}\right)} \equiv 1 \bmod p_i^{r_i} \mathcal{O}_{\mathbb{Q}[\sqrt{D}]} \text{ for } 1 \leq i \leq s.$$

With Theorem 2.27 and Proposition 2.28, the last congruence admits

$$d = \gcd\left(n - \left(\frac{D}{n}\right), p_i^{r_i - 1}(p_i - \left(\frac{D}{p_i}\right))\right) \tag{2.11}$$

$$= \gcd(n - \left(\frac{D}{n}\right), p_i - \left(\frac{D}{p_i}\right)) \tag{2.12}$$

many solutions. Among these solutions $\tau$, it is convenient to withdraw those for which $\tau - 1$ is not invertible modulo $p_i$. We will show that the only solution affected by this withdrawal is 1. Let us first note that

$$\begin{cases} \tau^{n-\left(\frac{D}{n}\right)} \equiv 1 \\ \tau^{p_i^{r_i-1}(p_i - \left(\frac{D}{p_i}\right))} \equiv 1 \end{cases} \implies \tau^d \equiv 1 \implies \tau^{p_i - \left(\frac{D}{p_i}\right)} \equiv 1 \bmod p_i^{r_i} \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$$

Let $\mathfrak{p}$ be a prime ideal containing $p_i \mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$. For an integer $k \geq 1$, we have

$$\begin{aligned} \tau \equiv 1 \bmod \mathfrak{p}^k &\implies \tau^{p_i} \equiv 1 \bmod \mathfrak{p}^{k+1} \\ &\implies 1 \equiv \tau^{p_i - \left(\frac{D}{p_i}\right)} \equiv \tau^{-\left(\frac{D}{p_i}\right)} \bmod \mathfrak{p}^{k+1} \\ &\implies \tau \equiv 1 \bmod \mathfrak{p}^{k+1}. \end{aligned}$$

Therefore, we have $\tau \equiv 1 \bmod \mathfrak{p}^{r_i}$. If $p_i$ is reducible in $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$, we have that $\tau \equiv 1 \bmod \overline{\mathfrak{p}}^{r_i}$, which implies that $\overline{\tau} \equiv \tau^{-1}$. In both cases, we have that $\tau \equiv 1 \bmod p_i^{r_i}$. The number of solutions of (2.11) is therefore $\gcd(n - (\frac{D}{n}), p - (\frac{D}{p_i})) - 1$. □

Now, similar to how Carmichael numbers completely defeat the weaker form of the Fermat test, we consider the numbers that completely defeat the weak Lucas test:

**Definition 2.32 (Lucas-Carmichael number)** *Let $D$ be a fixed integer. A Lucas-Carmichael number is a composite number $n$, relatively prime to $2D$, such that for all integers $P, Q$ with $\gcd(P, Q) = 1$, $D = P^2 - 4Q$ and $\gcd(n, QD) = 1$, $n$ is a $lpsp(P, Q)$.*
*Equivalently, for every $\tau$ that is a norm-1 element in $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$ where $\tau - 1$ is a unit in $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}/n$, we have $\tau^{n-(\frac{D}{n})} \equiv 1 \bmod n$.*

Williams showed in [26] the analogous theorem to Carmichael numbers.

**Theorem 2.33** *Let $D$ be a fixed integer, then $n$ is a $lpsp(P, Q)$ if and only if $n$ is square-free and $p_i - (\frac{D}{p_i}) \mid n - (\frac{D}{n})$ for every prime $p_i \mid n$.*

The question of the existence of an infinite number of Carmichael-Lucas numbers with respect to a fixed $D$ is still an open question. It should be noted that if $n$ is a Carmichael-Lucas number with respect to either $D = 1$ or $D$ a perfect square, then it is a Carmichael number. Thus, any result in this direction would be a generalization of the result concerning Carmichael numbers in [25], which in itself took 84 years to prove.

### 2.4.3 The Strong Lucas Probable Prime Test

In analogy to strong pseudoprimes, we define a stronger variant of the Lucas probable prime test, which leads to strong Lucas pseudoprimes. Let $P, Q \in \mathbb{Z}$ such that $D = P^2 - 4Q$, where $D$ is not a perfect square and $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$ be the ring of integers of $\mathbb{Q}[\sqrt{D}]$. Let $\alpha, \beta \in$ be the roots of the polynomial $X^2 - PX + Q$ in $\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}$ and for an integer $n \in \mathbb{N}$, let $(U_n)_{n\in\mathbb{N}}$ and $(V_n)_{n\in\mathbb{N}}$ denote the Lucas sequences defined as in (2.3). Finally, for an integer $n \in \mathbb{N}$, we denote $\varepsilon(n)$ the Jacobi Symbol $(\frac{D}{n})$, when $D$ is fixed.

**Theorem 2.34** *Let $P$ and $Q$ be integers and $D = P^2 - 4Q$. Let $p$ be a prime number not dividing $2QD$. Put $p - \varepsilon(p) = 2^\kappa q$ with $q$ odd. Then one of the following is satisfied:*

$$\begin{aligned} &p \mid U_q \\ &\textit{or} \\ &\textit{there exists } i \textit{ such that } 0 \le i < \kappa \textit{ and } p \mid V_{2^i q}. \end{aligned} \tag{2.13}$$

As with the other tests, we also have pseudoprimes for this test.

**Definition 2.35 (Strong Lucas Pseudoprimes)** *A composite number $n$ relatively prime to $2QD$ which satisfies (2.13) is called* a strong Lucas pseudoprime with respect to the parameters $P$ and $Q$. *For short we write $slpsp(P, Q)$.*

Usually we want $\varepsilon(n) = -1$, because otherwise the strong Lucas test is not independent from the Miller-Rabin test. We will discuss its reason in Subsection 2.4.4. The next theorem is from [4]. It says how many pairs $(P, Q)$ with $0 \leq P, Q < n$, $\gcd(Q, n) = 1$, $P^2 - 4Q \equiv D \bmod n$ exist, such that $n$ is a $slpsp(P, Q)$.

**Theorem 2.36** *[Theorem 1.5, [4]] Let $D$ be an integer and $p_1^{r_1} \dots p_s^{r_s}$ be the prime decomposition of an integer $n \geq 2$ relatively prime to $2D$. Put*

$$\begin{cases} n - \varepsilon(n) = 2^\kappa q \\ p_i - \varepsilon(p_i) = 2^{k_i} q_i \text{ for } 1 \leq i \leq s \end{cases} \quad \text{with } q, q_i \text{ odd},$$

*ordering the $p_i$'s such that $k_1 \leq \dots \leq k_s$. The number of pairs $(P, Q)$ with $0 \leq P, Q < n$, $\gcd(Q, n) = 1$, $P^2 - 4Q \equiv D \bmod n$ and such that $n$ is an* slpsp$(P, Q)$ *is expressed by the formula*

$$SL(D, n) = \prod_{i=1}^{s} (\gcd(q, q_i) - 1) + \sum_{j=0}^{k_1 - 1} 2^{js} \prod_{i=1}^{s} \gcd(q, q_i). \tag{2.14}$$

Using the fact that each Lucas sequence is in a one-to-one correspondence with the norm-1 elements $\tau$ in $\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}$ where $\tau - 1$ is a unit in $\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/(n)$, we get the following result:

**Lemma 2.37** *Let $n > 1$ be an integer relatively prime to $QD$ and let $\tau = \alpha\beta^{-1}$ in the ring $\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}$ ($\tau$ is well-defined as $Q = \alpha\beta$ is relatively prime to $n$). For $k \in \mathbb{N}$, we have the equivalences*

$$n \mid U_k \Leftrightarrow \tau^k = 1$$
$$n \mid V_k \Leftrightarrow \tau^k = -1.$$

*In particular, if $n$ is composite and relatively prime to $2QD$, it is a $slpsp(P, Q)$ if and only if*

$$\tau^q \equiv 1 \bmod n$$

*or*

$$\text{there exists } i \text{ such that } 0 \leq i < \kappa \text{ and } \tau^{2^i q} \equiv -1 \bmod n$$

*where $n - \varepsilon(n) = 2^\kappa q$ with $q$ odd.*

**Proof** With the Binet formula (2.14) we have

$$\begin{aligned} n \mid V_k &\Leftrightarrow V_k \in n\mathscr{O}_{\mathbb{Q}[\sqrt{D}]} \text{ since } n\mathscr{O}_{\mathbb{Q}[\sqrt{D}]} \cap \mathbb{Z} = n\mathbb{Z} \text{ and } V_k \in \mathbb{Z} \\ &\Leftrightarrow \alpha^k + \beta^k = \beta^k(1 + \tau^k) \in n\mathscr{O}_{\mathbb{Q}[\sqrt{D}]} \\ &\Leftrightarrow \tau^k + 1 \in n\mathscr{O}_{\mathbb{Q}[\sqrt{D}]}. \end{aligned}$$

The last equivalence holds as $\alpha\beta = Q$ and $1 = \gcd(Q, n) = \gcd(\alpha\beta, n)$, which holds if and only if $\gcd(\alpha, n) = \gcd(\beta, n) = 1$. Thus $\beta^k \notin n\mathbb{Z}$. For the other equivalence:

$$\begin{aligned} n \mid U_k &\Leftrightarrow U_k \in \mathcal{O}_{\mathbb{Q}[\sqrt{D}]} \\ &\Leftrightarrow \frac{\alpha^k - \beta^k}{\alpha - \beta} \in n\mathcal{O}_{\mathbb{Q}[\sqrt{D}]} \\ &\Leftrightarrow \frac{\beta^k(\tau^k - 1)}{\beta(\tau - 1)} = \frac{\beta^{k-1}(\tau^k - 1)}{\tau - 1} \in n\mathcal{O}_{\mathbb{Q}[\sqrt{D}]} \\ &\Leftrightarrow \tau^k - 1 \in n\mathcal{O}_{\mathbb{Q}[\sqrt{D}]}. \end{aligned}$$

□

### 2.4.4 The Square Root Problem

Although all theorems are true when $(\frac{D}{n}) = 1$, it is best to avoid this case, as then Fermat/ Miller-Rabin test and the (strong) Lucas test are not independent. Next we consider what is called the square root problem. Suppose that $P$ and $Q$ lead to $D$ such that $(\frac{D}{n}) = 1$. This does not mean that $D$ is a square modulo $n$, but it does increase the likelihood that that is so; for if $D$ is a square modulo $n$, then it is a square modulo each prime divisor of $n$ and so $(\frac{D}{n}) = 1$. We now analyze why is squareness bad.

#### When D is a perfect square

If $D$ is a non-zero perfect square, the Lucas test reduces to the Fermat test: Now our ring $\mathbb{Q}[\sqrt{D}]$ becomes $\mathbb{Q}$ and the ring of integers of $\mathbb{Q}$ becomes $\mathbb{Z}$, with this also our roots $\alpha, \beta$ of the polynomial $X^2 - PX + Q$ will be in $\mathbb{Z}$. We let $T = \alpha\beta^{-1} \bmod n$. Therefore from Lemma 2.37, we have

$$n \mid U_{n-1} \Leftrightarrow T^{n-1} \equiv 1 \bmod n,$$

which is just an ordinary Fermat test.

If $D$ is a non-zero perfect square the strong Lucas test reduces to the Rabin-Miller test: If $\gcd(n, 2D) = 1$, we can put $T = \alpha\beta^{-1} \bmod n$, where again $\alpha, \beta \in \mathbb{Z}$. Now from lemma 2.37 we get the following equivalences for $k \in \mathbb{N}$:

$$\begin{aligned} n \mid U_k &\Leftrightarrow T^k \equiv 1 \bmod n \\ n \mid V_k &\Leftrightarrow T^k \equiv -1 \bmod n. \end{aligned}$$

As then for all $i$ decompositions $n - 1 = n - \varepsilon(n) = 2^\kappa q$, $p_i - 1 = p_i - \varepsilon(p_i) = 2^{k_i} q_i$ are the same, we get that $n$ is a $slpsp(P, Q)$ if and only if it is a $spsp(P, Q)$.

However, there is an easy way to make sure that our chosen $D$ is not a perfect square. We can perform a test for squareness using Newton's method for square roots.

**Algorithm 3:** Checking for a perfect square: Newton's method

**Result:** Checks if an $n$-bit integer $D$ is a perfect square, i.e. (if $x^2 = D$ has an integer solution.)

Set $m = \lceil \frac{n}{2} \rceil$ and $i = 0$;

Select random $x_0$ s.t. $2^m > x_0 \geq 2^{m-1}$

**while** $x_i^2 \geq 2^m + D$ **do**
  $i = i + 1$;
  $x_i = \frac{1}{2}(x_{i-1} + \frac{c}{x_i})$

**if** $D = \lfloor {x_i}^2 \rfloor$ **then**
  status = perfect square

**else**
  status = not a perfect square

**Return** status

#### When D is a square modulo n

Now let $D$ not necessarily be a perfect square, but a square modulo $n$, which means that $(\frac{D}{n}) = 1$. We now establish a lemma that connects ordinary pseudoprimes and Lucas pseudoprimes.

**Lemma 2.38** *Let $n$ be an odd integer which is a $psp(b)$ and $psp(c)$. Then, whenever $P \equiv b + c$ and $Q \equiv bc \bmod n$, $n$ is a $lpsp(P, Q)$.*

**Proof** Let $\alpha$ and $\beta$ be the distinct roots of the polynomial $X^2 - PX + Q$. Then $\{\alpha, \beta\} \equiv \{b, c\} \bmod n$, because the quadratic polynomials $X^2 - PX + Q$ and $X^2 - (b + c)X + bc$ have coefficients that are congruent modulo $n$. Since $(\frac{D}{n}) = (\frac{P^2 - 4Q}{n}) = (\frac{(b-c)^2}{n}) = 1$, we get $U_{n-1} = \frac{b^{n-1} - c^{n-1}}{b-c} \equiv 1$. This shows that $n$ is a $lpsp(P, Q)$. □

Now, if $P$ and $Q$ are such that $D = r^2 \bmod n$, then the pair of simultaneous equations $P = b + c$ and $Q = bc$ can be easily solved modulo $n$ for $b$ and $c$ to get $b = \frac{(P-r)(n+1)}{2}$ and $c = \frac{(P+r)(n+1)}{2}$. If, for example, $n$ is a $psp(2)$, then it might well be a $psp(b)$ and a $psp(c)$ (assuming that $\gcd(n, bc) = 1$), because it might be a Carmichael number, or one of $b$ and $c$ might be $\pm 1$. This would mean, by Lemma 2.38 that $n$ will be a $lpsp(P, Q)$. This is bad because it means that the Lucas probable prime test will not be independent of the ordinary Fermat test.

#### When the Jacobi symbol is 1

Most of the following ideas are from [18]. Let $n = \prod p_i^{r_i}$ be the prime decomposition, $D$ not necessarily be a square modulo $n$, but $(\frac{D}{n}) = 1$. We count the number of distinct values $P$ modulo $n$ for which there is a $Q$ such that $n$ is a $lpsp(P,Q)$. By Theorem 2.31 it is $\prod((n-1, p_i \pm 1) - 1)$, where $n = \prod p_i^{r_i}$, and the choice of $\pm 1$ depends on $(\frac{D}{p_i})$ and $r_i$. Likewise $n$ is a $psp(a)$ for $\prod(n-1, p_i - 1)$ distinct values of $a$ modulo $n$. Now, the product of the gcds $(n-1, p_i - 1)$ and $(n-1, p_i + 1)$ is less than $2(p_1 + 1)$, but whenever $(\frac{D}{p_i}) = +1$, then the gcds are equal, so that both can be large. Thus, in many cases we would expect that if $n$ is a $lpsp(P,Q)$ for many values of $P$ with $(\frac{D}{n}) = +1$, then $n$ might also be a $psp(a)$ for many values for $a$. The computer calculations bear this out. See [18].

Now we do a similar analysis for the strong Lucas test. Again, $n = \prod p_i^{r_i}$ be the prime decomposition, $D$ not necessarily be a square modulo $n$, but $(\frac{D}{n}) = 1$. Let $n - 1 = 2^\kappa q$, $p_i - (\frac{D}{p_i}) = 2^{k_i} q_i$ and $p_i - 1 = 2^{l_i} s_i$ with $q_i, s_i$ odd. The number of pairs $(P,Q)$ such $0 \le P, Q < n$ with $D = P^2 - 4Q$, $\gcd(Q,n) = 1$ and such that $n$ is a $slpsp(P,Q)$ is $SL(D,n) = \prod_{i=1}^{s}(\gcd(q,q_i) - 1) + \sum_{j=0}^{k_1-1} 2^{js} \prod_{i=1}^{s} \gcd(q, p_i - 1)$. The number of bases $a$ such that $n$ is a $spsp(a)$ is equal to $S(n) = (1 + \sum_{j=0}^{k_1-1} 2^{js}) \prod_{i=1}^{s} \gcd(q, p_i - 1)$. Again the product of the gcds $(q, p_i - 1) \cdot (q, p_i + 1) \le$ cannot exceed $2(q, p_i + 1)$, but whenever $(\frac{D}{p_i}) = 1$, the gcds are the same, so they both can be large.

### 2.4.5 The Baillie-PSW Test

Baillie, Selfridge and Wagstaff [18] proposed in 1980 a probabilistic primality test that has become known as the Baillie-PSW test. Its power to find composites lies in combining a single Miller-Rabin test with base 2 with a (strong) Lucas test. The idea is that the two tests might be orthogonal to each other and thus it is very unlikely that a number $n$ will pass both parts. That is, $n$ being a probable prime of the first type does not affect the probability of $n$ being a probable prime of the second type, thus if $n$ passes both tests, we can be more certain that it is prime than if it merely passes several Miller-Rabin tests, or several Lucas tests. No odd composite integers $n$ have been reported to pass this combination of primality test if the parameters are chosen in an appropriate way.

**The Baillie-PSW Test**

Let $n$ be our odd integer, which we want to test for primality. We are going to declare it as a probable prime number if and only if it passes each of the following points:

1. If $n$ is divisible by any prime less than some convenient limit, for example 1000, then $n$ is composite.
2. If $n$ is not a (strong) pseudoprime base 2, then $n$ is composite.
3. Check if $n$ is not a perfect square and determine a pair of integers $(P, Q)$ through one of the following methods:
   - *Method A:* Let $D$ be the first element of the sequence $5, -7, 9, -11, 13, \ldots$ for which $(\frac{D}{n}) = -1$.
     Let $P = 1$ and $Q = (1 - D)/4$
   - *Method B:* Let $D$ be the first element of the sequence $5, 9, 13, 17, 21, \ldots$ for which $(\frac{D}{n}) = -1$.
     Let $P = \min\{m \in \mathbb{N} \mid m \text{ odd and } m > \sqrt{D}\}$ and $Q = (P^2 - D)/4$.
4. If $n$ is not a (strong) Lucas pseudoprime for our choice of parameters $P, Q$, then $n$ is composite. Otherwise, $n$ is a probable prime.

**Lemma 2.39** *Let $n = p_i m_i$ be an odd integer. Then the following holds*

$$\gcd(n-1, p_i - 1) = \gcd(n-1, m_i - 1)$$
$$\gcd(n+1, p_i + 1) = \gcd(n+1, m_i - 1).$$

**Proof** With $n - 1 = p_i m_i - 1 = (p_i - 1)m_i + (m_i - 1) = (m_i - 1)p_i + (p_i - 1)$, we get

$$\begin{aligned}\gcd(n-1, p_i - 1) &= \gcd((m_i - 1)p_i + (p_i - 1), p_i - 1)\\ &= \gcd((m_i - 1)p_i, p_i - 1) = \gcd(m_i - 1, p_i - 1)\\ \gcd(n-1, m_i - 1) &= \gcd((p_i - 1)m_i + (m_i - 1), m_i - 1)\\ &= \gcd((p_i - 1)m_i, m_i - 1) = \gcd(p_i - 1, m_i - 1),\end{aligned}$$

where the second last equality follows from the fact that for all $a, b$ we have $\gcd(a + b, a) = \gcd(a, b)$ and the last equality holds as $\gcd(a - 1, a) = 1$. For the second equality we use a similar argument. □

*Heuristic Argument:* The most interesting thing about the Lucas test is that if we choose the parameters $D$, $P$ and $Q$ as described in the second method, then the first 50 Carmichael numbers and several other base-2 Fermat pseudoprimes will never be Lucas pseudoprimes, see [18].

We give a heuristic argument why composite numbers rarely pass the Baillie-PSW test:

Let $P, Q$ and $D$ be integers such that $P^2 - 4Q = D$ and choose $n = p_1 \dots p_2$ with $\gcd(n, QD)$ and $(\frac{D}{n}) = 1$ such that it is both a $psp(b)$ and $lpsp(P, Q)$. Fermat's Little Theorem implies:

$$\begin{cases} b^{\gcd(n-1, p_i - 1)} \equiv 1 \text{ in } \mathbb{Z}/n\mathbb{Z} \\ \tau^{\gcd(n+1, p_i - \varepsilon(p_i))} \equiv 1 \text{ in } \mathscr{O}_{\mathbb{Q}[\sqrt{D}]}/n\mathscr{O}_{\mathbb{Q}[\sqrt{D}]} \end{cases} \quad \forall i,$$

where $\tau$ is the element associated with the pair $(P, Q)$ in Proposition 2.28. We let

$$\begin{cases} d_i = \gcd(n-1, p_i - 1) \\ d_i' = \gcd(n+1, p_i - \varepsilon(p_i)). \end{cases}$$

We have the equivalences

$$b^{n-1} \equiv 1 \bmod p_i \Leftrightarrow b \text{ is a } (p_i - 1)/d_i \text{th root } \bmod p$$
$$\tau^{n+1} = 1 \text{ in } \mathscr{O}/p_i\mathscr{O} \Leftrightarrow \tau \text{ is a } (p_i - \varepsilon(p_i))/d_i' \text{th root in } (\mathscr{O}/p_i\mathscr{O})^\wedge.$$

Heuristically, these relations have a very small chance of being true when the integers $d_i$ and $d_i'$ are small compared to the order of the groups $(\mathbb{Z}/p_i\mathbb{Z})^\times$ and $(\mathscr{O}/p_i\mathscr{O})^\wedge$.

For $\varepsilon(p_i) = 1$, we have

$$d_i d_i' = \gcd(n-1, p_i - 1)\gcd(n+1, p_1 - 1) \leq 2(p_1 - 1),$$

therefore it is not possible that both gcds are large.

For $\varepsilon(p_i) = -1$, we let $n = p_i m_i$. By Lemma 2.39, we have

$$\begin{cases} \gcd(n-1, p_i - 1) = \gcd(n-1, p_i - 1) = \gcd(n-1, m_i - 1) \\ \gcd(n+1, p_i + 1) = \gcd(n+1, p_i + 1) = \gcd(n+1, m_i - 1). \end{cases}$$

We conclude that

$$d_i d_i' = \gcd(n-1, p_i - 1)\gcd(n+1, p_i + 1) \leq 2(m_i - 1),$$

again by the same argument as above the gcds cannot be large, making $d_i$ and $d_i'$ rather small, which results in few pseudoprimes.

This "orthogonality" leads to the general belief that a combination of a Miller-Rabin and a (strong) Lucas test with properly chosen parameters could in fact be a deterministic primality test. Gilchrist [9] even confirmed that there are no Baillie-PSW pseudoprimes less than $2^{64}$, when using Method A. To date no composites have been found to pass such a combined test, it is therefore reasonable to conjecture that:

**Conjecture 2.40** *If $n > 1$ is a positive integer which can pass the combination of a strong pseudoprimality test and a Lucas test, then n is prime.*

Pomerance, Selfridge and Wagstaff [7] issued two challenges for an example of a composite number which passes both a strong pseudoprimality test base 2 and a Lucas test ($620), and/or a proof that no such number exists ($620). At the moment, the prizes are unclaimed; no counter-example has been found. Yet there is no proof that they cannot exist and in fact, Pomerance gave a heuristic argument in [17] that there are infinitely many Baillie-PSW pseudoprimes. The construction of a single example is a significant open problem in number theory.

Chapter 3

# An Analog to the Rabin-Monier Theorem for Lucas Pseudoprimes

## 3.1 The analog using $\varphi_D$

Rabin showed in [4] that there is an analog of Theorem 2.10 (the Rabin-Monier theorem) for strong Lucas pseudoprimes. In this chapter we will go through the details.

Recall that the Rabin-Monier theorem states that for all $n > 9$, we have $S(n)/\varphi(n) < 1/4$. However, the Euler phi function $\varphi(n)$ cannot be directly applied to the Lucas case. Thus, we will need to define a number theoretic function, which is a variant of $\varphi(n)$.

**Definition 3.1 (The $\varphi_D$ function)** *Let $D$ be an integer and for an integer $n$ let $\varepsilon(n)$ denote the Jacobi symbol $(\frac{D}{n})$. We introduce the following number theoretic function, which is defined only on integers relatively prime to $2D$:*

$$\begin{cases} \varphi_D(p^r) = p^{r-1}(p - \varepsilon(p)) \text{ for any prime } p \nmid 2D \text{ and } r \in \mathbb{N} \\ \varphi_D(p_1 \cdot p_2) = \varphi_D(p_1) \cdot \varphi_D(p_2) \text{ if } \gcd(p_1, p_2) = 1. \end{cases}$$

**Theorem 3.2 (The Rabin-Monier theorem for Lucas pseudoprimes)** *If $n$ is an odd composite integer not of the form $n = (2^{k_1}q_1 - 1)(2^{k_1}q_1 + 1)$, where both factors are prime, $q_1, k_1 \in \mathbb{Z}^+$ and $q_1$ odd, then*

$$SL(D, n) \leq \varphi_D(n)/4$$

*Also, the following inequality is always true: $SL(D, n) \leq \varphi_D(n)/2$.*

When $n$ is of the form $n = (2^{k_1}q_1 - 1)(2^{k_1}q_1 + 1)$, where both factors are prime and $k_1 \neq 1$, we will see that $SL(D, n) \leq \varphi_D(n)/2$ is quite an overestimate, Subsection 3.1.1 focuses on this special case.

In order the prove Theorem 3.2 we will need the following lemma, which is in analogy to Lemma 2.12.

**Lemma 3.3** *Let $p_1^{r_1} \dots p_s^{r_s}$ be the prime decomposition of an integer $n$ relatively prime to $2D$. Using the notation of Theorem (2.36), we have the inequalities:*

$$\frac{SL(D,n)}{\varphi_D(n)} \leq \begin{cases} \frac{1}{2^{s-1}} \prod_{i=1}^{s} \frac{\gcd(q,q_i)}{q_i}, \\ \frac{1}{2^{s-1}} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}}, \\ \frac{1}{2^{s-1+\delta_2+\dots+\delta_s}}, \text{ where } \delta_i = k_i - k_1. \end{cases}$$

The proof of Lemma 3.3 can be found in [4]. Equipped with this lemma, we are ready to prove Theorem 3.2.

**Proof (Theorem 3.2)** Let $s = 1$. From the second inequality of Lemma 3.3 it follows that $\frac{SL(D,n)}{\varphi_D(n)} \leq \frac{1}{p_1^{r_1-1}}$. Let $n \neq 9$, then either $p_1 \geq 5$ and $r_1 \geq 2$ or $p_1 \geq 3$ and $r_1 \geq 3$. Therefore, $SL(D,n)/\varphi_D(n) \leq 1/4$ follows directly. If $n = 9$, it is easy to verify using Theorem 2.36 that

$$p_1 - \varepsilon(p_1) = \begin{cases} 2 & \Rightarrow k_1 = 1, q_1 = 1 \quad \text{when } \varepsilon(p_1) = 1 \\ 4 & \Rightarrow k_1 = 2, q_1 = 1 \quad \text{when } \varepsilon(p_1) = -1. \end{cases}$$

In both cases $\gcd(q, q_1) = 1$. For the first case we get that $SL(D,n) = (1-1) + 2^0 \cdot 1 = 1$ and $\varphi_D(n) = 3(3-1) = 6$, and for the second case we get that $SL(D,n) = (1-1) + 2^0 \cdot 1 + 2^1 \cdot 1 = 3$ and $\varphi_D(n) = 3(3+1) = 12$. Thus $SL(D,n)/\varphi_D(n) \leq 1/4$ for both cases.

Now let $s = 2$. Lemma 3.3 yields the inequalities:

$$\frac{SL(D,n)}{\varphi_D(n)} \leq \begin{cases} 1/6 & \text{if } r_i \geq 2 \text{ for at least one } i \\ 1/4 & \text{if } \delta_2 = k_2 - k_1 \geq 1. \end{cases}$$

Thus we are only left with the case where $r_1 = r_2 = 1$ and $k_2 - k_1 < 1$, meaning that $k_2 = k_1$.

Let us suppose that $q_1 \neq q_2$. The first inequality of Lemma 3.3 yields

$$\frac{SL(D,n)}{\varphi_D(n)} \leq \frac{1}{2} \frac{\gcd(q,q_1)}{q_1} \frac{\gcd(q,q_2)}{q_2}.$$

We will show by contradiction that at least one of $\gcd(q,q_i)/q_i \leq 1/3$ . For this, let's suppose both $\frac{\gcd(q,q_1)}{q_1}, \frac{\gcd(q,q_2)}{q_2} > 1/3$. Since all $q, q_1, q_2$ are odd, we must have that $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$, meaning that both $q_1 \mid q$ and $q_2 \mid q$. Therefore, both $q_1$ and $q_2$ must divide

$$\begin{aligned} 2^\kappa q =& p_1 p_2 - \varepsilon(p_1 p_2) \\ =& (2^{k_1} q_1 + \varepsilon(p_1))(2^{k_1+\delta_2} q_2 + \varepsilon(p_2)) - \varepsilon(p_1 p_2) \\ =& 2^{2k_1+\delta_2} q_1 q_2 + 2^{k_1} q_1 \varepsilon(p_2) + 2^{k_1+\delta_2} q_2 \varepsilon(p_1) + \varepsilon(p_1)\varepsilon(p_2) - \varepsilon(p_1 p_2) \\ =& 2^{2k_1+\delta_2} q_1 q_2 + 2^k (q_1 \varepsilon(p_2) + 2^{\delta_2} q_2 \varepsilon(p_1)) \\ =& 2^{2k_1+\delta_2} q_1 q_2 \pm 2^k (q_1 \pm 2^{\delta_2} q_2). \end{aligned}$$

Therefore, we have that $q_1 \mid q_2$ and $q_2 \mid q_1$, which is only possible when $q_1 = q_2$, which is contradictory to our hypothesis that $q_1 \neq q_2$.

Now let's suppose that $r_1 = r_2 = 1$, $k_1 = k_2$ and $q_1 = q_2$. Then $p_1 - \varepsilon(p_1) = 2^{k_1}q_1$ and $p_2 - \varepsilon(p_2) = 2^{k_1}q_1$. If $\varepsilon(p_1) = \varepsilon(p_2)$, we would have that $p_1 = p_2$, which contradicts our hypothesis that $n$ is a product of two distinct primes. Thus, we must have that $\varepsilon(p_1) = -\varepsilon(p_2)$. Without loss of generality we assume that $\varepsilon(p_1) = -1$. Thus, $n = (2^{k_1}q_1 + 1)(2^{k_1}q_1 - 1)$ and $\varphi_D(n) = (2^{k_1}q_1)^2$, which is the exception, where, using the first inequality of Lemma 3.3, we have that $\frac{SL(D,n)}{\varphi_D(n)} \leq \frac{1}{2} \cdot (\frac{gcd(q,q_1)}{q_1})^2 \leq \frac{1}{2}$.

Now let $s \geq 3$. The theorem holds trivially for $s = 3$ using either of the first two inequalities of Lemma 3.3, which finally(!) completes our proof. □

### 3.1.1 Analysis of the case $n = (2^{k_1}q_1 - 1)(2^{k_1}q_1 + 1)$

The Rabin-Monier Theorem for the Lucas test says that $\frac{SL(D,n)}{\varphi_D(n)} \leq 1/4$ for all odd composite $n \neq 9$ except when $n$ is of the form $n = (2^{k_1}q_1 - 1)(2^{k_1}q_1 + 1)$, where both factors are prime. Now let us analyse the case where $n$ is of the latter form. Thus, for this subsection, we suppose that $n$ is always of the form $n = (2^{k_1}q_1 - 1)(2^{k_1}q_1 + 1)$, with both factors prime.

**Corollary 3.4** *For the decomposition $n - \varepsilon(n) = 2^\kappa q$ with $q$ odd, we have that $\kappa = 2k_1$ and $q = q_1^2$, and therefore, $n - \varepsilon(n) = n + 1 = 4^{k_1}q_1^2$.*

**Proof** With

$$\begin{aligned} p_1 &= 2^{k_1}q_1 - 1, \quad p_1 - \varepsilon(p_1) = 2^{k_1}q_1 \quad \text{where } \varepsilon(p_1) = -1 \\ p_2 &= 2^{k_1}q_1 + 1, \quad p_2 - \varepsilon(p_2) = 2^{k_1}q_1 \quad \text{where } \varepsilon(p_2) = 1, \end{aligned}$$

we get that $\varepsilon(n) = \varepsilon(p_1)\varepsilon(p_2) = -1$, and

$$\begin{aligned} &n = p_1p_2 = (2^{k_1}q_1 - 1)(2^{k_1}q_1 + 1) = 4^{k_1} \cdot q_1^2 - 1 \stackrel{!}{=} 2^\kappa \cdot q - 1 \\ &\Rightarrow \kappa = 2k_1, q = q_1^2. \end{aligned}$$

Thus, $n - \varepsilon(n) = 4^{k_1} \cdot q_1^2$ and $\gcd(q, q_1) = \gcd(q_1^2, q_1) = q_1$. □

**Corollary 3.5** *For $n$ of the above form, we have*

$$SL(D,n) = (q_1 - 1)^2 + \frac{4^{k_1} - 1}{3} \cdot q_1^2$$

*and*

$$\frac{SL(D,n)}{\varphi_D(n)} = \frac{(q_1 - 1)^2 + \frac{4^{k_1}-1}{3} \cdot q_1^2}{4^{k_1} \cdot q_1^2}.$$

**Proof** Using the fact that $\sum_{j=0}^{k_1-1} 4^j = \frac{4^{k_1}-1}{3}$, as the sum is a geometric series and applying Theorem 2.36, we have that

$$SL(D,n) = \prod_{i=1}^{q}((q,q_i)-1) + \sum_{j=0}^{k_1} 2^{j\cdot 2} \prod_{i=1}^{2}(q,q_i) = (q_1-1)^2 + \frac{4^{k_1}-1}{3}\cdot q_1^2.$$

Also for $\varphi_D(n)$ we have $\varphi_D(n) = (p_1 - \varepsilon(p_1))\cdot(p_2-\varepsilon(p_2)) = 4^{k_1}\cdot q_1^2$. Thus

$$\alpha_D(n) = \frac{SL(D,n)}{\varphi_D(n)} = \frac{(q_1-1)^2 + \frac{4^{k_1}-1}{3}\cdot q_1^2}{4^{k_1}\cdot q_1^2} \qquad \square$$

The only case when $\alpha_D(n) \leq 1/4$ is for $k_1 = q_1 = 1$, as then $\alpha_D(n) = 1/4$.

**Lemma 3.6**

$$\frac{SL(D,n)}{\varphi_D(n)} < \frac{1}{2} \text{ for all } n \text{ of the form } n = (2^{k_1}q_1 - 1)(2^{k_1}q_1+1),$$

*where both factors are prime,* $k_1, q_1 \in \mathbb{N}$ *and* $q_1$ *odd.*

**Proof** By Lemma 3.5 we have

$$\begin{aligned}\frac{SL(D,n)}{\varphi_D(n)} &= \frac{(q_1-1)^2 + \frac{4^{k_1}-1}{3}\cdot q_1^2}{4^{k_1}\cdot q_1^2} \\ &< \frac{q_1^2 + \frac{4^{k_1}-1}{3}\cdot q_1^2}{4^{k_1}\cdot q_1^2} = \frac{1+\frac{4^{k_1}-1}{3}}{4^{k_1}} = \frac{2}{3\cdot 4^{k_1}} + \frac{1}{3}.\end{aligned}$$

Since $1/4^{k_1}$ is a decreasing function in $k_1$, we can upper bound it by using $k_1 = 1$. Thus we get

$$\frac{SL(D,n)}{\varphi_D(n)} = \frac{2}{3\cdot 4^{k_1}} + \frac{1}{3} \leq \frac{1}{6} + \frac{1}{3} = \frac{1}{2}. \qquad \square$$

Unless $k_1 = 1$, this is quite an overestimate. However when we fix $k_1 = 1$, we see that the larger $q_1$ gets, the closer this ratio gets to $1/2$.

**Lemma 3.7** *Let* $k_1 = 1$*, then we have*

$$\lim_{q_1\to\infty} \frac{SL(D,n)}{\varphi_D(n)} = \frac{1}{2}.$$

**Proof**

$$\lim_{q_1\to\infty} \frac{SL(D,n)}{\varphi_D(n)} = \lim_{q_1\to\infty} \frac{(q_1-1)^2 + \frac{4^{k_1}-1}{3}\cdot q_1^2}{4^{k_1}\cdot q_1^2} = \lim_{q_1\to\infty} \frac{1}{6} - \frac{1}{2q_1} + \frac{1}{4q_1^2} + \frac{1}{3} = \frac{1}{2}.$$

$\square$

When $k_1$ is kept fixed, the ratio $\alpha_D(n)$ converges to a specific number as $q_1$ increases.

**Lemma 3.8** *Let $k_1 \in \mathbb{N}$ be fixed. Then*

$$\lim_{q_1 \to \infty} \frac{SL(D,n)}{\varphi_D(n)} = \lim_{q_1 \to \infty} \frac{1}{6 \cdot 4^{k_1 - 1}} + \frac{1}{3}.$$

**Example 3.9** *(a) For $k_1 = 2$ then $\lim_{q_1 \to \infty} \frac{SL(D,n)}{\varphi_D(n)} = \frac{9}{24} = 0.375$,*

*(b) when $k_1 = 3$, then $\lim_{q_1 \to \infty} \frac{SL(D,n)}{\varphi_D(n)} = \frac{33}{96} = 0.34375$,*

*(c) $k_1 = 19$ then $\lim_{q_1 \to \infty} \frac{SL(D,n)}{\varphi_D(n)} = \frac{45812984491}{137438953472} \approx 0.3333333333357$.*

It seems to be the case as as $k_1$ increases, the ratio gets closer to $1/3$.

We have now considered what happens when $k_1$ is a fixed positive integer. However our ratio $SL(D,n)/\varphi_D(n)$ depends on both $k_1$ and $q_1$. Now we fix $q_1$, and consider what happens when $k_1$ increases.

**Lemma 3.10** *Let $q_1 \in \mathbb{N}$ be odd and fixed, then*

$$\lim_{k_1 \to \infty} \frac{SL(D,n)}{\varphi_D(n)} = \frac{1}{3}$$

**Proof**

$$\lim_{k_1 \to \infty} \frac{3(q_1-1)^2 - q_1^2}{3 \cdot 4^{k_1} \cdot q_1^2} + \frac{1}{3} = \lim_{k_1 \to \infty} \frac{1}{4^{k_1}} - \frac{2}{4^{k_1} q_1} + \frac{1}{4^{k_1} q_1^2} - \frac{1}{3 \cdot 4^{k_1}} + \frac{1}{3} = \frac{1}{3}. \quad \square$$

We can show however, that even though the ratio converges to $1/3$, it is almost always bigger than $1/3$.

**Lemma 3.11** *For all $q_1, k_1 \in \mathbb{N}$ with $q_1$ odd, except when $q_1 = 1$, we have*

$$\frac{SL(D,n)}{\varphi_D(n)} > \frac{1}{3}.$$

**Proof**

$$\frac{SL(D,n)}{\varphi_D(n)} = \frac{3(q_1-1)^2 - q_1^2}{3 \cdot 4^{k_1} \cdot q_1} + \frac{1}{3}$$

Since $3 \cdot 4^{k_1} \cdot q_1 > 0$ for all $k_1, q_1 \in \mathbb{N}$, it remains to show that $3(q_1-1)^2 - q_1^2 > 0$ for all $q_1 \neq 1$. Solving this quadratic equation, we get that $3(q_1 - 1)^2 - q_1^2 \leq 0$ if and only if $-(\frac{3}{2} \pm \frac{\sqrt{3}}{2}) \leq q_1 \leq \frac{3}{2} \pm \frac{\sqrt{3}}{2}$, meaning the only possible value in our case, since $q_1$ is an odd positive integer is $q_1 = 1$. $\square$

When $q_1 = 1$, then $\frac{SL(D,n)}{\varphi_D(n)} = \frac{1}{3} - \frac{1}{3 \cdot 4^{k_1}} < \frac{1}{3}$.

## 3.2 An analog to the Rabin-Monier Theorem using $n$

For the Miller-Rabin test we can directly conclude that $S(n) < n/4$, which means that at least $3/4$-th of the bases $1 \leq a \leq n-1$ are witnesses for the compositeness of $n$. However, as we will see in Lemma 4.10, $\varphi_D(n)$ is not bounded by $n$. Therefore, the result of Theorem 3.2 cannot be directly translated like the Rabin-Monier theorem for the Miller-Rabin test. Nevertheless, Arnault showed in [4] the following more powerful result:

**Theorem 3.12** *Let $D$ be an integer and $n$ a composite number relaitvely prime to $2D$ and distinct from 9. For every integer $D$, we have*

$$SL(D,n) \leq \frac{4n}{15},$$

*except if $n$ is the product if $n = (2^{k_1}q_1 - 1)(2^{k_1}q_1 + 1)$ of twin primes with $q_1$ odd and such that the Legendre symbols satisfy $\varepsilon(2^{k_1}q_1 - 1) = -1$, $\varepsilon(2^{k_1}q_1 + 1) = 1$, where we have $SL(D,n) \leq n/2$.*

Chapter 4

# Average case error estimate for the strong Lucas probable prime test

We have already talked about the average case error estimate of the Miller-Rabin test and its relation to the Rabin-Monier theorem. We have seen that the former does not directly follow from the knowledge of the latter. However, Damgård et al. [8] showed bounds for the average case probability, which are stated in Theorem 2.13. As such average error estimates do not exist for the strong Lucas test, we are interested in establishing such results. We know by Theorem 3.12 that $SL(D,n) \leq 4n/15$ for every odd integer $n$, which is not a product of twin primes. In this chapter, we make a thorough analysis of this error probability, and establish results similar to Theorem 2.13. We obtain these bounds by closely following the methods used in [8].

Recall the definition of $SL(D,n)$. Let $\alpha_D(n) = \frac{SL(D,n)}{\varphi_D(n)}$ for $n > 1$, where $n$ is odd. Thus $\alpha_D(n) \leq 1/4$ for odd composites $n \neq (2^{k_1}q_1 - 1)\cdot(2^{k_1}q_1 + 1)$, where both factors are prime. Let $n - \varepsilon(n) = 2^\kappa q$, with $q$ odd. Also let $n = p_1^{r_1} \dots p_s^{r_s}$ be the prime decomposition of an integer relatively prime to $2D$, ordering the $p_i's$ such that $k_1 \leq \cdots \leq k_s$ in the decomposition $p_i - \varepsilon(p_i) = 2^{k_i}q_i$, where $q_i$ is odd. This implies that $k_1$ is the largest integer such that $2^{k_1} \mid p_i - \varepsilon(p_i)$ for all $i = 1, 2, \dots, s$.

Let $\omega(n)$ denote the number of distinct prime factors of $n$ and let $\Omega(n)$ denote the number of prime factors of $n$ counted with multiplicity. Here we have that $\omega(n) = s$ and $\Omega(n) = \sum_{i=1}^{s} r_i$. We shall always let $p$ denote a prime number. We begin by first stating lemmas which will be used in later proofs.

**Lemma 4.1** *In the factorization $n - \varepsilon(n) = 2^\kappa q$ and $p_i - \varepsilon(p_i) = 2^{k_i}q_i$, we have that $2^{k_i} \mid 2^\kappa$, which implies $k_i \leq \kappa$ for all $i$.*

**Proof** Since the Jacobi symbol is completely multiplicative, the following equation holds:

$$\begin{aligned}2^\kappa \cdot q =& n - \varepsilon(n) = \prod_{i=1}^{s} p_i^{r_i} - \varepsilon(\prod_i p_i^{r_i}) = \prod_{i=1}^{s}\Big[\sum_{j=0}^{r_i}(2^{k_i}q_i)^{r_i-j}\cdot\varepsilon(p_i)^j\Big] - \prod_{i=1}^{s}\varepsilon(p_i)^{r_i}\\ =&\prod_{i=1}^{s}\Big[((2^{k_i}q_i)^{r_i} + (2^{k_i}q_i)^{r_i-1}\cdot\varepsilon(p_i) + \cdots + (2^{k_i}q_i)^2\cdot\varepsilon(p_i)^{r_i-2}\\ &+ (2^{k_i}q_i)\cdot\varepsilon(p_i)^{r_i-1} + \varepsilon(p_i)^{r_i}\Big] - \prod_{i=1}^{s}\varepsilon(p_i)^{r_i}.\end{aligned}$$

The term $\prod_{i=1}^{s}\varepsilon(p_i)^{r_i}$ completely cancels out, thus in every term the factor $2^{k_i}$ appears at least once, and can thus be factored out:

$$\begin{aligned}2^\kappa q =& \prod_{i=1}^{s} 2^{k_i}\Big[((2^{k_i r_i-1})\cdot q_i^{r_i} + (2^{k_i(r_i-1)-1})\cdot q_i^{r_i-1}\cdot\varepsilon(p_i) + \ldots\\ &+ (2^{2k_i-1}\cdot q_i^2\cdot\varepsilon(p_i)^{r_i-2} + 2^{k_i-1}\cdot q_i\cdot\varepsilon(p_i)^{r_i-1}\Big]\\ =& 2^{\sum_{i=1}^{s}k_i}\prod_{i=1}^{s}\Big[((2^{k_i r_i-1})\cdot q_i^{r_i} + (2^{k_i(r_i-1)-1})\cdot q_i^{r_i-1}\cdot\varepsilon(p_i) + \ldots\\ &+ 2^{2k_i-1}\cdot q_i^2\cdot\varepsilon(p_i)^{r_i-2} + 2^{k_i-1}\cdot q_i\cdot\varepsilon(p_i)^{r_i-1}\Big].\end{aligned}$$

Thus, we see that $2^{k_i} \mid 2^\kappa$ for all $i$, therefore $k_i \leq \kappa$. □

**Lemma 4.2** *Let $k_1, j, s \in \mathbb{N}$. Then*

$$\Big(1 + \sum_{j=0}^{k_1-1} 2^{j\cdot s}\Big) \leq 2\cdot 2^{(k_1-1)\cdot s}.$$

**Proof** This may be proved by showing that for all $m \in \mathbb{N}$ and fixed $s \in \mathbb{N}$ we have that $1 + \sum_{j=0}^{m} 2^{j\cdot s} \leq 2\cdot 2^{m\cdot s}$. We do this by induction on $m$:
*Base case:* $m = 1$ : As $2 \leq 2^s$ we have $1 + \sum_{j=0}^{1} 2^{j\cdot s} = 1 + 1 + 2^s \leq 2^s + 2^s = 2\cdot 2^s$. So our claim holds for $m = 1$.
*Inductive hypothesis:* Suppose the theorem holds for all $m \leq m_0$.
*Inductive step:* Then as $s \geq 1$

$$\begin{aligned}1 + \sum_{j=0}^{m_0+1} 2^{j\cdot s} &= 1 + \sum_{j=0}^{m_0} 2^{j\cdot s} + 2^{(m_0+1)\cdot s} \leq 2\cdot 2^{m_0\cdot s} + 2^{(m_0+1)\cdot s}\\ &\leq 2^s\cdot 2^{m_0\cdot s} + 2^s\cdot 2^{m_0\cdot s} = 2\cdot 2^{s\cdot(m_0+1)},\end{aligned}$$

which is exactly what we wanted to show. □

**Lemma 4.3** *If $n = p_1^{r_1}\ldots p_s^{r_s} > 1$ is odd, then*

$$\alpha_D(n) \leq 2^{1-s}\prod_{i=1}^{s} p^{1-r_i}\frac{(p-\varepsilon(p), n-\varepsilon(n))}{p-\varepsilon(p)} \leq 2^{1-\Omega(n)}\prod_{i=1}^{s}\frac{(p-\varepsilon(p), n-\varepsilon(n))}{p-\varepsilon(p)}.$$

**Proof** We see that the identity $\sum_{i=1}^{s}(r_i-1)=\Omega(n)-s$ trivially holds. Thus, $2^{(1-s)}=2^{1-\Omega(n)+\sum_{i=1}^{s}(r_i-1)}=2^{1-\Omega(n)}\prod_{i=1}^{s}2^{r_i-1}$. Using the fact that $\frac{2}{p}\leq 1$ for every prime $p$ and $r_i\geq 1$ for all $i$, the second inequality follows by

$$2^{1-s}\prod_{i=1}^{s}p^{1-r_i}=2^{1-\Omega(n)}\prod_{i=1}^{s}\frac{2^{r_i-1}}{p^{r_i-1}}\leq 2^{1-\Omega(n)}\prod_{i=1}^{s}(\frac{2}{p})^{r_i-1}\leq 2^{1-\Omega(n)}.$$

For the first inequality we use

$$SL(D,n)=\prod_{i=1}^{s}((q,q_i)-1)+\sum_{j=0}^{k_1-1}2^{js}\prod_{i=1}^{s}(q,q_i)\leq\Big(1+\sum_{j=0}^{k_1-1}2^{js}\Big)\prod_{i=1}^{s}(q,q_i).$$

Using this upper bound and the definition of $\varphi_D(n)$, we get

$$\begin{aligned}\alpha_D(n)=\frac{SL(D,n)}{\varphi_D(n)}&\leq\Big(1+\sum_{j=0}^{k_1-1}2^{j\cdot s}\Big)\cdot\prod_{i=1}^{s}\frac{(q_i,q)}{p_i^{r_i-1}(p_i-\varepsilon(p_i))}\\&=\Big(1+\sum_{j=0}^{k_1-1}2^{j\cdot s}\Big)\prod_{i=1}^{s}\frac{(p_i-\varepsilon(p_i),q)}{p_i^{r_i-1}(p_i-\varepsilon(p_i))}.\end{aligned}$$

For two coprime numbers $a,b\in\mathbb{N}$ and for all $c\in\mathbb{N}$ we have $(c,a\cdot b)=(c,a)\cdot(c,b)$. In the factorization $n-\varepsilon(n)=2^{\kappa}\cdot q$, the two factors $2^{\kappa}$ and $q$ are coprime, thus we get

$$\prod_{i=1}^{s}(p_i-\varepsilon(p_i),n-\varepsilon(n))=\prod_{i=1}^{s}(p_i-\varepsilon(p_i),q)\cdot(p_i-\varepsilon(p_i),2^k)=$$
$$\prod_{i=1}^{s}(p_i-\varepsilon(p_i),q)\cdot(2^{k_i}q_i,2^k)=\prod_{i=1}^{s}(p_i-\varepsilon(p_i),q)\cdot(2^{k_i},2^k).$$

By Lemma 4.1 we know that $k_i\leq\kappa$ for all $i=1,\ldots,s$, and using the way we have defined the order of the $p_i$'s, i.e. $k_1\leq\cdots\leq k_s$, we have

$$\prod_{i=1}^{s}(p_i-\varepsilon(p_i),q)(2^{k_i},2^{\kappa})=\prod_{i=1}^{s}2^{k_i}\cdot(p_i-\varepsilon(p_i),q)\geq 2^{k_1\cdot s}\prod_{i=1}^{s}(p_i-\varepsilon(p_i),q).$$

$$\begin{aligned}\Rightarrow\prod_{i=1}^{s}(p_i-\varepsilon(p_i),q)&\leq 2^{-k_1\cdot s}\prod_{i=1}^{s}(p_i-\varepsilon(p_i),n-\varepsilon(n))\\&=2^{-sk_1}\prod_{i=1}^{s}(p_i-\varepsilon(p_i),n-\varepsilon(n)).\end{aligned}$$

By Lemma 4.2 we know that

$$1+\sum_{j=0}^{k_1-1}2^{j\cdot s}\leq 2\cdot 2^{(k_1-1)\cdot s}.$$

Therefore,

$$\begin{aligned}\alpha_D(n) = \frac{SL(D,n)}{\varphi_D(n)} &\leq \Big(1 + \sum_{j=0}^{k_1-1} 2^{j\cdot s}\Big) \cdot \prod_{i=1}^{s} \frac{(p_i - \varepsilon(p_i), q)}{p_i^{r_i-1} \cdot (p_i - \varepsilon(p_i))} \\ &\leq \Big(1 + \sum_{j=0}^{k_1-1} 2^{j\cdot s}\Big) 2^{-k_1\cdot s} \prod_{i=1}^{s} \frac{(p_i - \varepsilon(p_i), n - \varepsilon(n))}{p_i^{r_i-1} \cdot (p_i - \varepsilon(p_i))} \\ &\leq 2 \cdot 2^{(k_1-1)\cdot s} \cdot 2^{-k_1\cdot s} \prod_{i=1}^{s} \frac{(p_i - \varepsilon(p_i), n - \varepsilon(n))}{p_i^{r_i-1} \cdot (p_i - \varepsilon(p_i))} \\ &= 2^{1-s} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}} \cdot \frac{(p_i - \varepsilon(p_i), n - \varepsilon(n))}{p_i - \varepsilon(p_i)},\end{aligned}$$

which proves the assertion. □

**Remark 4.4** We see that the formula for $SL(D,n)$ looks *similar* to the formula for $S(n)$. However, we must keep in mind that for Lucas test we consider the factorization of $p_i - \varepsilon(p_i) = 2^{k_i} q_i$, whereas for the Miller-Rabin test we consider the factorization $p_i - 1 = 2^{k_i} q_i$, so depending on whether $\varepsilon(p_i) = -1, 1$, the $k_i$'s and $q_i$'s could be completely different in the factorization.

The following lemmas and corollaries are used in a later proof.

**Lemma 4.5** *If $t \in \mathbb{R}$ with $t \geq 1$, then*

$$\sum_{n=\lfloor t \rfloor + 1}^{\infty} \frac{1}{n(n-1)} = \frac{1}{\lfloor t \rfloor} < \frac{2}{t}.$$

**Proof**

$$\sum_{n=\lfloor t \rfloor + 1}^{\infty} \frac{1}{n(n-1)} = \lim_{k\to\infty} \sum_{n=\lfloor t \rfloor + 1}^{k} \frac{1}{n-1} - \frac{1}{n} = \lim_{k\to\infty} \frac{1}{\lfloor t \rfloor} + \frac{1}{k} = \frac{1}{\lfloor t \rfloor} < \frac{2}{t},$$

where we used the partial fractial decomposition of $\frac{1}{n(n-1)}$ and the fact that $\sum_n (\frac{1}{n-1} - \frac{1}{n})$ is a telescope sum. □

**Corollary 4.6** *Let $k, t \in \mathbb{N}$, then*

$$2^{-2\sqrt{t(k-1)}} < 2^{\sqrt{\frac{t}{k-1}} - 2\sqrt{tk}}. \tag{4.1}$$

**Proof** Since both sides of the inequality $\sqrt{tk} < \sqrt{t(k-1)} + \sqrt{\frac{t}{4(k-1)}}$ are positive, we get that squaring them preserves the inequality, yielding $tk < tk + \frac{t}{4(k-1)}$. This is trivially true when $t, k \geq 1$. □

**Corollary 4.7** *For all $j, k \in \mathbb{N}$, we have*

$$jt + \frac{k-1}{j} \geq 2\sqrt{t(k-1)}.$$

**Proof** With $0 \leq (j\sqrt{t} - \sqrt{k-1})^2 = j^2 t - 2\sqrt{t(k-1)} + (k-1)$ the corollary directly follows for $j, k > 0$. □

## 4.1 A simple estimate

Let $C_m = \{n \in \mathbb{N} : n \text{ odd, composite and } \alpha_D(n) > 2^{-m}\}$. Let $M_k$ denote the set of odd $k$-bit integers. For $k \geq 2$, we have $|M_k| = 2^{k-2}$. We are concerned with the proportion in $M_k$ of those odd integers which are also in $C_m$.

**Theorem 4.8** *If $m, k$ are positive integers with $m + 1 \leq 2\sqrt{k-1}$, then*

$$\frac{|C_m \cap M_k|}{|M_k|} < 8 \sum_{j=2}^{m} 2^{m-j-\frac{k-1}{j}}.$$

**Proof** From Lemma 4.3, $n \in C_m$ implies $\Omega(n) \leq m$, as

$$2^{-m} < \alpha_D(n) \leq 2^{1-\Omega(n)} \prod_{p|n} \frac{(p - \varepsilon(p), n - \varepsilon(n))}{p - \varepsilon(p)} \leq 2^{1-\Omega(n)} \Rightarrow m \geq \Omega(n).$$

Now let $N(m, k, j) = \{n \in C_m \cap M_k : \Omega(n) = j\}$. Thus

$$|C_m \cap M_k| = \sum_{j=2}^{m} |N(m, k, j)|.$$

Suppose $n \in N(m, k, j)$, where $2 \leq j \leq m$. Let $p$ denote the largest prime factor of $n$. Since $2^{k-1} < n < 2^k$, we have $p > 2^{(k-1)/j}$. Let $d_D(p, n) = \frac{p - \varepsilon(p)}{(p - \varepsilon(p), n - \varepsilon(n))}$. From Lemma 4.3 and the definition of $C_m$, we have

$$2^m > \frac{1}{\alpha_{D(n)}} \geq 2^{\Omega(n)-1} d_D(p, n) = 2^{j-1} d_D(p, n),$$

so that $d_D(p, n) < 2^{m+1-j}$.
Given $p, d$, where $p$ is a prime with the property that $p > 2^{(k-1)/j}$ and $d$ is such that $d \mid p - \varepsilon(p)$ and $d < 2^{m+1-j}$, we want to get an upper bound for the number $n \in N(m, k, j)$ with the largest prime factor $p$ such that and $d_D(p, n) = d$. Let $S_{k,d,p} = \{n \in M_k : p \mid n, d = \frac{p - \varepsilon(p)}{(p - \varepsilon(p), n - \varepsilon(n))}, n \text{ composite}\}$. The size of the set $S_{k,d,p}$ is at most the number of solutions of the system

$$n \equiv 0 \bmod p, \quad n \equiv \pm 1 \bmod \tfrac{p - \varepsilon(p)}{d}, \quad p < n < 2^k$$

i.e. at most the set $R_{k,d,p} = \{n \in \mathbb{Z} : n \equiv 0 \bmod p, n \equiv \pm 1 \bmod \frac{p-\varepsilon(p)}{d}, p < n < 2^k\}$, via the Chinese Remainder Theorem $R_{k,d,p}$ has less than $\frac{2^k d}{p(p-\varepsilon(p))}$ elements.

Let us look at the parity of $(p-\varepsilon(p))/d$. Let $q_p$ be the odd part of the decomposition $p-\varepsilon(p)$ and let $\nu_2(p) = \max\{2^i : 2^i \mid p,\ i \in \mathbb{N}\}$. We have that

$$\frac{p-\varepsilon(p)}{d} = (p-\varepsilon(p), n-\varepsilon(n)) = 2^{\nu_2(p)} \cdot l, \text{ where } l \text{ is odd.}$$
$$= 2^{\nu_2(p)} \frac{q_p}{q_p} = p - \varepsilon(p) \frac{l}{q_p}, \text{ where } \frac{l}{q_p} \text{ is odd.}$$

Therefore, $(p-\varepsilon(p))/d$ is even.

If $S_{k,d,p} \neq \emptyset$, then there exists an $n \in S_{k,d,p}$ with $(n-\varepsilon(n), p-\varepsilon(p)) = (p-\varepsilon(p))/d$, and thus $(p-\varepsilon(p))/d$. Thus, we only need to consider $d$ and $p$, such that $(p-\varepsilon(p))/d$ is even. We conclude that

$$|N(m,k,j)| \leq \sum_{p > 2^{(k-1)/j}} \sum_{\substack{d \mid p-\varepsilon(p) \\ d < 2^{m+1-j} \\ (p-\varepsilon(p))/d \in 2\mathbb{Z}}} \frac{2^k d}{p(p-\varepsilon(p))}$$
$$= 2^k \sum_{d < 2^{m+1-j}} \sum_{\substack{p > 2^{(k-1)/j} \\ d \mid p-\varepsilon(p) \\ (p-\varepsilon(p))/d \in 2\mathbb{Z}}} \frac{d}{p(p-\varepsilon(p))}.$$

Now, for the inner sum we have,

$$\sum_{\substack{p > 2^{(k-1)/j} \\ d \mid p-\varepsilon(p) \\ \frac{p-\varepsilon(p)}{d} \in 2\mathbb{Z}}} \frac{d}{p(p-\varepsilon(p))} < \sum_{2ud > 2^{\frac{k-1}{j}} - \varepsilon(p)} \frac{d}{(2ud+\varepsilon(p))2ud}$$
$$= \frac{1}{4d} \sum_{2ud > 2^{\frac{k-1}{j}} - \varepsilon(p)} \frac{1}{(u + \frac{\varepsilon(p)}{2d})u} \leq \frac{1}{4d} \sum_{2ud > 2^{\frac{k-1}{j}} - \varepsilon(p)} \frac{1}{u(u - \frac{1}{2d})}$$
$$\leq \frac{1}{4d} \sum_{u > \frac{2^{\frac{k-1}{j}} - \varepsilon(p)}{2d}} \frac{1}{u(u-1)} < \frac{1}{4d} \frac{2}{\frac{2^{\frac{k-1}{j}} - \varepsilon(p)}{2d}} = \frac{1}{2^{\frac{k-1}{j}} - \varepsilon(p)},$$

where the last inequality follows from Lemma 4. Using this estimate, we get

$$|N(m,k,j)| \leq 2^k \sum_{d < 2^{m+1-j}} \frac{1}{2^{\frac{k-1}{j}} - \varepsilon(p)} = 2^k \cdot \frac{2^{m+1-j} - 1}{2^{\frac{k-1}{j}} - \varepsilon(p)}.$$

Using the fact that $\forall j,k \geq 1$, we have that $j^2 - 2\sqrt{k-1}j + (k-1) = (j - \sqrt{k-1})^2 \geq 0$, we get the inequality

$$j + \frac{k-1}{j} \geq 2\sqrt{k-1}.$$

Using this inequality and our hypothesis that $m+1 \leq 2\sqrt{k-1}$, we have $m+1 \leq j + (k-1)/j$. Thus

$$\frac{2^{m+1-j}-1}{2^{\frac{k-1}{j}} - \varepsilon(p)} \leq \frac{2^{m+1-j}-1}{2^{\frac{k-1}{j}} - 1} \leq \frac{2^{m+1-j}}{2^{\frac{k-1}{j}}} = 2 \cdot 2^{m-j-\frac{k-1}{j}}.$$

The last inequality is true because for $m+1-j \leq (k-1)/j$ we have

$$\begin{aligned} & 2^{m+1-j+\frac{k-1}{j}} - 2^{\frac{k-1}{j}} \leq 2^{m+1-j+\frac{k-1}{j}} - 2^{m+1-j} \\ \Leftrightarrow & (2^{m+1-j}-1)2^{\frac{k-1}{j}} \leq 2^{m+1-j}(2^{\frac{k-1}{j}} - 1) \\ \Leftrightarrow & \frac{2^{m+1-j}-1}{2^{\frac{k-1}{j}} - 1} \leq \frac{2^{m+1-j}}{2^{\frac{k-1}{j}}}. \end{aligned}$$

Therefore, $N(m,k,j)| \leq 2 \cdot 2^{k+m-j-\frac{k-1}{j}}$. Combining everything and using the fact that $|M_k| = 2^{k-2}$ yields

$$\frac{|C_m \cap M_k|}{|M_k|} = \frac{\sum_{j=2}^{m} |N(m,k,j)|}{2^{k-2}} \leq 8 \sum_{j=2}^{m} 2^{m-j-\frac{k-1}{j}}. \qquad \square$$

## 4.2 The average case error probability

Let $D$ be fixed. We are trying to find a numerical upper bound for the probability that a number chosen uniformly at random from the set of $k$-bit integers is composite given that it passes $t$ independent iterations of the strong Lucas test with randomly chosen bases $(P,Q)$.

Let $\overline{\alpha}_D(n) = \frac{SL(D,n)}{n}$ be the fraction of elements in $\{1,2,\ldots,n\}$ for which the strong Lucas probable prime test is positive. This makes sense, as $SL(D,n)$ is the number of pairs $(P,Q)$, where $1 \leq P \leq n$ that satisfies some properties, and we already know that for every $P$ there exists exactly one $Q$, such that $P = D^2 - 4Q$. Therefore, $SL(D,n)$ actually only counts those $P$'s chosen from the set $1,\ldots,n$.

Let $X$ denote the event that $n$ is composite and let $Z_t$ denote the event that $n$ is chosen uniformly at random from $M_k$ and that it has passed $t$ consecutive rounds of the strong Lucas test with uniformly chosen bases $(P, Q(P,D))$.

Also let $\pi(x)$ denote the prime counting function up to $x$ and let $\sum'$ denote the sum over composite integers. Using the law of conditional probability, we have for

$$q_{k,t} = \mathbb{P}[X \mid Z_t] = \frac{\mathbb{P}[X \cap Z_t]}{\mathbb{P}[Z_t]} = \frac{\left(\frac{\sum'_{n \in M_k} \overline{\alpha}_D(n)^t}{\sum'_{n \in M_k} 1}\right)}{\left(\frac{\sum_{n \in M_k} \overline{\alpha}_D(n)^t}{\sum_{n \in M_k} 1}\right)}$$
$$= \frac{\sum'_{n \in M_k} \overline{\alpha}_D(n)^t}{\sum_{n \in M_k} \overline{\alpha}_D(n)^t} \leq \frac{\sum'_{n \in M_k} \overline{\alpha}_D(n)^t}{\sum_{p \in M_k} \overline{\alpha}_D(p)^t} = \frac{\sum'_{n \in M_k} \overline{\alpha}_D(n)^t}{\pi(2^k) - \pi(2^{k-1})}, \tag{4.2}$$

where $p$ is prime.

In order to get an upper estimate for $q_{k,t}$, it will suffice to find an upper estimate for the final sum in equation (4.2) and a lower estimate for $\pi(2^k) - \pi(2^{k-1})$.

The next proposition can be found in [8] as Proposition 2.

**Proposition 4.9** *For k an integer at least 21, we have*

$$\pi(2^k) - \pi(2^{k-1}) > (0.71867)\frac{2^k}{k}. \tag{4.3}$$

We already know a lot about $\alpha_D(n)$, but for our probability $q_{k,t}$ we will need $\overline{\alpha}_D(n)$. In this thesis we look at two different approaches of bounding $\overline{\alpha}_D(n)$ using $\alpha_D(n)$. One approach holds for the most general case, whereas the other approach uses trial division by small primes. Naturally, the one not making assumptions will yield a weaker estimate, which is still good enough for small $t$. However, we will see that for large $t$ this estimate is useless. Assuming that $n$ is not divisible by small primes does in many cases not impose high restrictions, as many cryptolibraries use trial division prior to expensive primality tests in order to speed up prime generation and primality testing.

We start to estimate $\varphi_D(n)$ as follows:

$$\varphi_D(n) = \prod_{i=1}^{s} p_i^{r_i-1}(p_i - \varepsilon(p_i)) \leq \prod_{i=1}^{s} p_i^{r_i-1}(p_i + 1) = \prod_{i=1}^{s}(p_i^{r_i} + p_i^{r_i-1}). \tag{4.4}$$

Then we will use our two different approaches. However, let us first investigate if this is not an overestimate, i.e. if integers $n$ exist such that $\varepsilon(p_i) = -1$ for all primes $p_i \mid n$.

**Lemma 4.10** *For $D = 2$, there exists infinitely many integers $n$, such that $\varphi_D(n) = \prod_{i=1}^{s}(p_i^{r_i} + p_i^{r_i-1})$, where $n = \prod_{i=1}^{s}(p_i^{r_i}$ is the prime decomposition.*

**Proof** For $D = 2$, we get by Gauss' law of quadratic reciprocity that

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1 \text{ if } p \equiv \pm 1 \mod 8 \\ -1 \text{ if } p \equiv \pm 3 \mod 8. \end{cases}$$

This is true as $p$ only depends on $p \mod 8$: if $p \equiv a \mod 8$, then $p = a + 8k, k \in \mathbb{N}$, $(\frac{2}{p}) = (-1)^{\frac{p^2-1}{8}} = (-1)^{\frac{a^2+16ak+64k^2-1}{8}} = (-1)^{\frac{a^2-1}{8}}$. Therefore we let $n$ to be a product of primes, such that every prime is congruent to $\pm 3$ mod 8; we have constructed an integer $n$ such that $(\frac{D}{p}) = -1$ for all primes $p \mid n$. Infinitely many such $n$'s exist, as by Dirichlet's theorem on arithmetic progressions, infinitely many primes $p \equiv \pm 3 \mod 8$ exist. □

So we see that the bound $\varphi_D(n) \leq \prod_{i=1}^{s} p_i^{r_i-1}(p_i + 1)$ is tight and can in general not be weakened.

## 4.3 Treatment of the numbers with large contribution in the analysis

By equality (4.2) our goal is to upper $\sum'_{n \in M_k} \overline{\alpha}_D(n)^t$. However, we do not know how to proceed using $\overline{\alpha}_D(n)$. As we will see, taking $\alpha_D(n)$ in the sum instead, will enable us get good estimations, thus, we must find a way to bound $\overline{\alpha}_D(n)$ by $\alpha_D(n)$, see Sections 4.4 and 4.5. Therefore, the numbers that add most to our probability are the ones with largest $\alpha_D(n)$ value. Since $\alpha_D(n) \leq 1/2$ for all $n$ by Theorem 3.12, we see that the set $C_1$ is empty. $C_2$ is not empty, we have already identified them as the set of integers $\{n = (2^{k_1}q_1 - 1)(2^{k_1}q_1 + 1) : \text{each factor is prime}, k_1, q_1 \in \mathbb{N}\}$. Thus, if $n < x$, we want to know how many tuples $(k_1, q_1) \in \mathbb{N}^2$ of such a form exist. If we can we find an upper bound for the number of such tuples, we can also upper bound $\mid C_2 \cap M_k \mid$, say $\mid C_2 \cap M_k \mid \leq f(k)$ for some function $f$ depending on $k$. Depending on if and how we treat those numbers, we get variants of our estimates:

1. We don't treat them separately.
2. We use an upper bound of the number of twin prime, which we will establish shortly, and treat $C_2$ differently in our analysis.
3. We count how many numbers of the form $m(m + 2)$ for some $m \in \mathbb{N}$ exist, as every product of twin primes is included in this set, and treat them differently in our analysis.
4. We sieve them at the beginning, making sure our number $n$ is not a twin prime, which enables us to exclude $C_2$ completely.

### 4.3.1 An upper bound of the number of twin primes

For variant 2 we want to count how many twin prime tuples up to $x$ exist. Let $n \in \mathbb{N}$ and let us recall the definition of $\Omega(n)$ in Section 4 counting the number of prime factors of $n$ counted with multiplicity. For any $x \in \mathbb{N}$, we define the twin prime counting function $\pi_2(x)$, which counts the number of twin-prime tuples up to $x$ as follows:

$$\pi_2(x) = \mid \{p \leq x : \Omega(p+2) = 1\} \mid .$$

It is conjectured that $\pi_2(x) \to \infty$ for $x \to \infty$.

Riesel and Vaughan [21] showed the following lemma.

**Lemma 4.11** *For $x > e^{42}$, we have*

$$\pi_2(x) < \frac{16\alpha x}{(7.5 + \log(x))\log(x)},$$

*where $\alpha$ is called the Twin Prime Constant,*

$$\alpha = \prod_{p>2} \left(1 - \frac{1}{(p-2)^2}\right) = \prod_{p>2} \frac{p(p-2)}{(p-1)^2} \approx 0.6602\ldots$$

Using their estimate we obtain the following lemma:

**Lemma 4.12** *For $k \geq 122$ we have*

$$\mid M_k \cap C_2 \mid < 20.3\frac{2^{k/2}}{k^2}.$$

**Proof** Since $n = p(p+2)$ is a $k$-bit integer, $p$ is a $k/2$-bit integer. Thus we are considering the number of twin primes up to $2^{k/2}$. We also have $\frac{16\alpha x}{(7.5+\log(x))\log(x)} < \frac{16\alpha x}{\log(x)^2}$. With Lemma 4.11 we obtain

$$\begin{aligned}\pi_2(2^{k/2}) &< \frac{16\alpha 2^{k/2}}{(\log(2^{k/2}))^2} = \frac{2^{k/2}}{(k/2)^2}\frac{16\alpha}{(\log_2(e))^2} < \frac{16 \cdot 4 \cdot 0.66017}{(\log_2(e))^2}\frac{2^{k/2}}{k^2} \\ &< 20.3\frac{2^{k/2}}{k^2}.\end{aligned}$$

for $2^{k/2} > e^{42}$, which implies that $k > \frac{84}{\log_e(2)} \approx 121.186$, thus $k \geq 122$ as $k \in \mathbb{N}$. Therefore, we get that

$$\mid M_k \cap C_2 \mid < 20.3\frac{2^{k/2}}{k^2}. \qquad \square$$

The next numbers with large $\alpha_D(n)$ values are the elements of $C_3, C_4$ and $C_5$. As we need them for future results, we bound $\mid M_k \cap C_m \mid$ for $m = 3, 4, 5$ by using Lemma 4.8.

**Lemma 4.13** *Let $k \geq 122$, then*

$$\begin{aligned} \mid M_k \cap C_3 \mid &\leq (2.52)2^{k-\frac{k}{3}} \\ \mid M_k \cap C_4 \mid &\leq (2.39)2^{k-\frac{k}{4}} \\ \mid M_k \cap C_5 \mid &\leq (2.37)2^{k-\frac{k}{5}}. \end{aligned}$$

**Proof** Since $m \leq 5$, we have that $m+1 \leq 2\sqrt{k-1}$ for $k \geq 10$, thus we may use Theorem 4.8. Therefore, we have that $\frac{|C_m \cap M_k|}{|M_k|} < 2^3 \sum_{j=2}^{m} 2^{m-j-\frac{k-1}{j}}$ and with this $\mid C_m \cap M_k \mid \leq 2^{k+1} \sum_{j=2}^{m} 2^{m-j-\frac{k-1}{j}}$. Let us bound each of those sums. We are looking for a $c_m \in \mathbb{R}$ such that $\mid M_k \cap C_m \mid \leq c_m 2^{k-\frac{k}{m}}$.

$$\begin{aligned} \mid M_k \cap C_3 \mid &\leq 2^{k+1} \sum_{j=2}^{3} 2^{3-j-\frac{k-1}{j}} = 2^{k+1}(2^{1-\frac{k-1}{2}} + 2^{-\frac{k-1}{3}}) \leq c_1 2^{k-\frac{k}{3}} \\ &\Rightarrow 2^{\frac{5}{2}-\frac{k}{6}} + 2^{\frac{4}{3}} \leq c_1. \\ &\text{With } k \geq 122 \text{ we get} \\ &\quad 2^{\frac{5}{2}-\frac{k}{6}} + 2^{\frac{4}{3}} \leq 2^{\frac{5}{2}-\frac{122}{6}} + 2^{\frac{4}{3}} \leq 2.52 \\ &\Rightarrow c_1 = 2.52. \\ \mid M_k \cap C_4 \mid &\leq 2^{k+1} \sum_{j=2}^{4} 2^{4-j-\frac{k-1}{j}} = 2^{k+1}(2^{2-\frac{k-1}{2}} + 2^{1-\frac{k-1}{3}} + 2^{-\frac{k-1}{4}}) \leq c_2 2^{k-\frac{k}{4}} \\ &\Rightarrow 2^{\frac{7}{2}-\frac{k}{4}} + 2^{\frac{7}{3}-\frac{k}{12}} + 2^{\frac{5}{4}} \leq c_2. \\ &\text{With } k \geq 122 \text{ we get} \\ &\quad 2^{\frac{7}{2}-\frac{k}{4}} + 2^{\frac{7}{3}-\frac{k}{12}} + 2^{\frac{5}{4}} \leq 2^{\frac{7}{2}-\frac{122}{4}} + 2^{\frac{7}{3}-\frac{122}{12}} + 2^{\frac{5}{4}} \leq c_2 \\ &\Rightarrow c_2 = 2.39. \\ \mid M_k \cap C_5 \mid &\leq 2^{k+1} \sum_{j=2}^{5} 2^{5-j-\frac{k-1}{j}} \\ &= 2^{k+1}(2^{3-\frac{k-1}{2}} + 2^{2-\frac{k-1}{3}} + 2^{1-\frac{k-1}{4}} + 2^{-\frac{k-1}{5}}) \leq c_3 2^{k-\frac{k}{5}} \\ &\Rightarrow 2^{\frac{9}{2}-\frac{3k}{10}} + 2^{\frac{10}{3}-\frac{2k}{15}} + 2^{\frac{9}{4}-\frac{k}{20}} + 2^{\frac{6}{5}} \leq c_3. \\ &\text{With } k \geq 122 \text{ we get} \\ &\quad 2^{\frac{9}{2}-\frac{3k}{10}} + 2^{\frac{10}{3}-\frac{2k}{15}} + 2^{\frac{9}{4}-\frac{k}{20}} + 2^{\frac{6}{5}} \leq 2.37 \\ &\Rightarrow c_3 = 2.37. \end{aligned}$$

□

#### Why counting twin-numbers does not yield an improvement

As discussed earlier we can upper bound the number of twin primes by upper bounding the number of twin numbers, namely numbers of the form

$n = m(m+2)$, where $m \in \mathbb{N}$.

**Corollary 4.14** *For $m \in \mathbb{N}$ odd, there are at most $\frac{\sqrt{x}}{2}$ such $n = m(m+2) \leq x$.*

*Also, we have $\mid C_2 \cap M_k \mid \leq \frac{1}{2}2^{k/2}$.*

**Proof** If $n = m(m+2) \leq x$, then $m^2 \leq x$. We use the fact that $m$ is odd, thus we have at most $\sqrt{x}/2$ such $n \leq x$. We immediately get that $\mid C_2 \cap M_k \mid \leq \frac{1}{2}2^{k/2}$ for any $k \geq 2$ and $x = 2^k$. □

We see that the estimate in Lemma 4.12 yields a tighter estimate, thus we discard the approach counting twin numbers using Corollary 4.14.

## 4.4 Bounding $q_{k,t}$

In this section, we establish implicit bounds for $q_{k,t}$ without making any assumptions on the $k$-bit integer tested for primality. Let us first prove the necessary results.

Akbary and Friggstad showed in [3] the following proposition.

**Proposition 4.15**

$$\frac{n}{\varphi(n)} \leq (1.07)e^{\gamma}\log(\log(n)) \qquad \text{for } n \geq 2^{78}.$$

Using their result we obtain an explicit upper bound for $\varphi_D$-function.

**Lemma 4.16** *For integers $k \geq 79$ and $n \in M_k$ we have*

$$\varphi_D(n) < n \cdot (1.07)e^{\gamma}\log(k),$$

*where $\gamma$ is the Euler-Mascheroni constant:*

$$\gamma = \lim_{n\to\infty}\Big(\sum_{k=1}^{n}\frac{1}{k} - \ln(n)\Big) < 0.58.$$

**Proof**

$$\varphi_D(n) = \prod_{i=1}^{s} p_i^{r_i-1}(p_i - \varepsilon(p_i)) \leq \prod_{i=1}^{s} p_i^{r_i-1}(p_i+1) = n\prod_{i=1}^{s}\frac{p_i+1}{p_i} = n\prod_{i=1}^{s}\Big(1+\frac{1}{p_i}\Big)$$
$$\leq n\prod_{i=1}^{s}\Big(1+\frac{1}{p_i-1}\Big) = n\prod_{i=1}^{s}\frac{p_i}{p_i-1} = n\prod_{i=1}^{s}\frac{1}{\frac{p_i-1}{p_i}} = n\prod_{i=1}^{s}\frac{1}{1-\frac{1}{p_i}} = \frac{n}{\prod_{i=1}^{s}\left(1-\frac{1}{p_i}\right)}. \tag{4.5}$$

The product $\prod_{p|n}\left(1-\frac{1}{p}\right)$ looks familiar: for the "normal" Euler phi function $\varphi(n)$ we have:

$$\varphi(n)=\prod_{i=1}^{s}p_i^{r_i-1}(p_i-1)=\prod_{i=1}^{s}p_i^{r_i}-p_i^{r_i-1}=\prod_{i=1}^{s}p_i^{r_i}\Big(1-\frac{1}{p_i}\Big)=n\prod_{i=1}^{s}\Big(1-\frac{1}{p_i}\Big).$$

Thus, we have

$$\prod_{i=1}^{s}(1-\frac{1}{p_i})=\frac{\varphi(n)}{n}.$$

Plugging this in (4.5), we get

$$\varphi_D(n)\leq n\cdot\frac{n}{\varphi(n)}.$$

By Proposition 4.15 we have for $k\geq 79, n\in M_k$

$$\begin{aligned}\varphi_D(n)&\leq n\cdot\frac{n}{\varphi(n)}<n\cdot(1.07)e^{\gamma}\log(\log(n))<n\cdot(1.07)e^{\gamma}\log(\log(2^k))\\&=n\cdot(1.07)e^{\gamma}\log(k\log(2))=n\cdot(1.07)e^{\gamma}(\log(k)+\log(\log(2)))\\&<n\cdot(1.07)e^{\gamma}(\log(k)-0.366)<n\cdot(1.07)e^{\gamma}\log(k),\end{aligned}$$

which finishes our proof. □

Therefore, we immediately get

**Corollary 4.17** *For $k\geq 79$ and $n\in M_k$ we have*

$$\overline{\alpha}_D(n)\leq(1.07)e^{\gamma}\log(k)\alpha_D(n).$$

### 4.4.1 First numerical results

Now that we have obtained an estimate for $\overline{\alpha}_D(n)$ using $\alpha_D(n)$, we can proceed with our analysis.

**Proposition 4.18** *For any integers $k, M, t$ with $3\leq M\leq 2\sqrt{k-1}-1, t\geq 1$ and $k\geq 79$ we have*

$$\sum{}'_{n\in M_k}\overline{\alpha}_D(n)^t\leq 2^{k-2+t(1-M)}\log^t(k)+2^{k+1+2t}\log^t(k)\sum_{j=2}^{M}\sum_{m=j}^{M}2^{m(1-t)-j-\frac{k-1}{j}}.$$

**Proof** Note that our hypothesis implies $k\geq 5$. We know that $C_1\cap M_k=\emptyset$. Thus by Corollary 4.17 we have

$$\begin{aligned}\sum{}'_{n\in M_k}\overline{\alpha}_D(n)^t&=\sum_{m=2}^{\infty}\sum_{n\in M_k\cap C_m\setminus C_{m-1}}\overline{\alpha}_D(n)^t\\&\leq\sum_{m=2}^{\infty}\sum_{n\in M_k\cap C_m\setminus C_{m-1}}((1.07)e^{\gamma}\log(k))^t\alpha_D(n)^t.\end{aligned}$$

We have that $n \in C_m \Leftrightarrow \alpha_D(n) > 2^{-m}$ and $n \notin C_{m-1} \Leftrightarrow \alpha_D(n) \leq 2^{-(m-1)}$. Thus, we get for $n \in C_m \setminus C_{m-1}$ that $2^{-m} < \alpha_D(n) \leq 2^{-(m-1)}$. Using this and the fact that $1.07e^{\gamma} < 2$ we get that

$$\sum_{m=2}^{\infty} \sum_{n \in M_k \cap C_m \setminus C_{m-1}} ((1.07)e^{\gamma} \log(k))^t \alpha_D(n)^t \tag{4.6}$$

$$\leq \sum_{m=2}^{\infty} ((1.07)e^{\gamma} \log(k))^t 2^{-(m-1)t} |M_k \cap C_m \setminus C_{m-1}|$$

$$< \sum_{m=2}^{\infty} \log^t(k) 2^{t-(m-1)t} |M_k \cap C_m \setminus C_{m-1}| \tag{4.7}$$

$$\leq \log^t(k) \Big( 2^{t-Mt} |M_k \setminus C_M| + \sum_{m=2}^{M} 2^{t-(m-1)t} |M_k \cap C_m| \Big) \tag{4.8}$$

$$= \log^t(k) \Big( 2^{t(1-M)} |M_k \setminus C_M| + \sum_{m=2}^{M} 2^{(2-m)t} |M_k \cap C_m| \Big). \tag{4.9}$$

Using Theorem 4.8 and the above estimate we have

$$\begin{aligned} \sum\nolimits'_{n \in M_k} \overline{\alpha}_D(n)^t &\leq \log^t(k) \Big( 2^{k-2+t(1-M)} + 2^{k+1+2t} \sum_{m=2}^{M} \sum_{j=2}^{m} 2^{m(1-t)-j-\frac{k-1}{j}} \Big) \\ &= \log^t(k) \Big( 2^{k-2+t(1-M)} + 2^{k+1+2t} \sum_{j=2}^{M} \sum_{m=j}^{M} 2^{m(1-t)-j-\frac{k-1}{j}} \Big). \end{aligned}$$

□

### 4.4.2 An estimate for $q_{k,1}$

We begin with the following inequality:

**Theorem 4.19** *For $k \geq 2$ we have $q_{k,1} < \log(k) k^2 4^{2.3-\sqrt{k}}$.*

**Proof** From (4.2) we have for $k \geq 21$ that

$$q_{k,1} \leq \frac{\sum\nolimits'_{n \in M_k} \overline{\alpha}_D(n)}{\pi(2^k) - \pi(2^{k-1})}. \tag{4.10}$$

Using Proposition 4.18 with $t = 1$ and $k \geq 79$ we get for any integer $M$ with $3 \leq M \leq 2\sqrt{k-1} - 1$

$$\begin{aligned} \sum\nolimits'_{n \in M_k} \overline{\alpha}_D(n) &\leq \log(k) \Big( 2^{k-1-M} + 2^{k+3} \sum_{j=2}^{M} 2^{-j-\frac{k-1}{j}} \sum_{m=j}^{M} 1 \Big) \\ &= \log(k) \Big( 2^{k-1-M} + 2^{k+3} \sum_{j=2}^{M} (M+1-j) 2^{-j-\frac{k-1}{j}} \Big). \end{aligned}$$

Using $t = 1$ in Corollary 4.7, we get $j + \frac{k-1}{j} \geq 2\sqrt{k-1}$ for all $j, k \geq 1$. Thus we have

$$\sum{}'_{n \in M_k} \overline{\alpha}_D(n) \leq \log(k)\Big(2^{k-1-M} + 2^{k+3-2\sqrt{k-1}} \sum_{j=2}^{M} (M+1-j)\Big). \tag{4.11}$$

Evaluating the sum $\sum_{j=2}^{M}(M+1-j) = M(M-1)\cdot 2^{-1}$, and letting $M = \lfloor 2\sqrt{k-1} - 1 \rfloor$, which implies that $M > 2\sqrt{k-1} - 2$ and $M \leq 2\sqrt{k-1} - 1$, we have

$$\begin{aligned}
\sum{}'_{n \in M_k} \overline{\alpha}_D(n) &\leq \log(k)\Big(2^{k-1-M} + 2^{k+2-2\sqrt{k-1}} M(M-1)\Big) \\
&< \log(k)\Big(2^{k+1-2\sqrt{k-1}}(1 + 2\cdot(4(k-1) - 6\sqrt{k-1} + 2)\Big) \\
&= \log(k)\Big(2^{k+1-2\sqrt{k-1}}(8k - 3 - 12\sqrt{k-1})\Big) \\
&< \log(k)k2^{k+4-2\sqrt{k-1}}.
\end{aligned} \tag{4.12}$$

By Corollary 4.6 with $t = 1$ we know that

$$2^{-2\sqrt{k-1}} < 2^{\frac{1}{\sqrt{k-1}} - 2\sqrt{k}}.$$

Using this we obtain from inequality (4.12) for $k \geq 100$

$$\sum{}'_{n \in M_k} \overline{\alpha}_D(n) < \log(k)k2^{4 + \frac{1}{\sqrt{99}} + k - 2\sqrt{k}}. \tag{4.13}$$

With $\frac{2^{4+\frac{1}{\sqrt{99}}}}{0.71867} < 4^{2.3}$ we get by Proposition 4.9 and inequality (4.13) for $k \geq 100$, that

$$q_{k,1} = \frac{\log(k)k^2 \cdot 2^{4+\frac{1}{\sqrt{99}} - 2\sqrt{k}}}{0.71867} < \log(k)k^2 4^{2.3-\sqrt{k}}.$$

But for $k \leq 100$ we have that $\log(k)k^2 4^{2.3-\sqrt{k}} > 1$ , so this upper bound is trivially true for $k \leq 100$. □

Damgård et al. established in [8] an average case error estimate for the Miller-Rabin test, namely they obtained that $p_{k,1} < k^2 4^{2-\sqrt{k}}$, where $p_{k,t}$ is the probability that an integer chosen uniformly at random from the set $M_k$ is composite given that it has passed $t$ consecutive round of the Miller-Rabin test. To get an idea of how good the bounds are, we let $k = 1024$ and get:

$$\begin{aligned}
&q_{1024,1} < 9.6 \cdot 10^{-12} \text{ using Thereom 4.19,} \\
&p_{1024,1} < 9.1 \cdot 10^{-13} \text{ using Theorem 2 in [8].}
\end{aligned}$$

### 4.4.3 An estimate for $q_{k,t}$

Now we consider the average case error estimate for a number, which has passed $t$ consecutive rounds of the strong Lucas test with respect to randomly chosen bases. We need the following lemma.

**Lemma 4.20** *Let $f(k,t) = \sqrt{\frac{k}{t}}\frac{2^{2t}}{1-2^{1-t}}$. For fixed $k_0 \geq 1$, $f(k_0,t)$ is a monotonically increasing function for all $t \geq 2$ and for a fixed $t_0 \geq$, $f(k,t_0)$ is a monotonically increasing function for all $k \geq 1$.*

**Proof** Let us first fix $k_0 \geq 1$. If $\frac{\partial f(k_0,t)}{\partial t} \geq 0$, then our function is monotonically increasing in $t$.

$$\begin{aligned}\frac{\partial f(k_0,t)}{\partial t} &= -\frac{\sqrt{\frac{k_0}{t}}2^{2t-1}}{(1-2^{1-t})t} + \frac{2^{2t+1}\ln 2\sqrt{\frac{k_0}{t}}}{1-2^{1-t}} - \frac{2^{t+1}\ln 2\sqrt{\frac{k_0}{t}}}{(1-2^{1-t})^2} \\ &= \frac{\sqrt{\frac{k_0}{t_0}}2^t}{1-2^{1-t}}\Big(-\frac{2^{t-1}}{t} + 2^{t+1}\ln(2) - \frac{2\ln(2)}{1-2^{1-t}}\Big) \geq 0.\end{aligned}$$

This holds if

$$-\frac{2^{t-1}}{t} + 2^{t+1}\ln(2) - \frac{2\ln(2)}{1-2^{1-t}} \geq 0 \tag{4.14}$$

and

$$\frac{\sqrt{\frac{k_0}{t_0}}2^t}{1-2^{1-t}} \geq 0. \tag{4.15}$$

(4.14) is equivalent to $2 \geq \frac{1}{2\ln(t)t} + \frac{1}{2^{t-1}-1}$, which is true for $t \geq 1.74322$. Since $t \in \mathbb{N}$, we have $\frac{\partial f(k_0,t)}{\partial t} \geq 0$ for $t \geq 2$. (4.15) holds for all $t > 1$. Thus, for fixed $t_0$, it follows directly that our function is monotonically increasing in $k$. □

**Theorem 4.21** *For $k,t$ integers with $k \geq 79$, $3 \leq t \leq k/9$ or $k \geq 88$, $t = 2$ we have*

$$q_{k,t} < k^{3/2}\frac{2^t}{\sqrt{t}}\log^t(k)4^{2.12-\sqrt{tk}}.$$

**Proof** Assume $k \geq 79$ and $t \geq 2$. When using Proposition 4.18, which says $\sum'_{n\in M_k}\overline{\alpha}_D(n)^t \leq \log^t(k)\Big(2^{k-2+t(1-M)} + 2^{k+1+2t}\sum_{j=2}^{M}\sum_{m=j}^{M}2^{m(1-t)-j-\frac{k-1}{j}}\Big)$. We want to estimate $\sum_{m=j}^{M}2^{m(1-t)}$. We do this by seeing that $\sum_{m=j}^{M}2^{m(1-t)} = \sum_{m=0}^{M}2^{m(1-t)} - \sum_{m=0}^{j-1}2^{m(1-t)} = \frac{2^{1-t}(2^j-2^M)}{1-2^{1-t}} \leq \frac{2^{j(1-t)}}{1-2^{1-t}}$ since $j \leq M$. Using this estimate in Proposition 4.18, we get that

$$\sum'_{n\in M_k}\overline{\alpha}_D(n)^t \leq 2^{k-2+t(1-M)}\log^t(k) + \frac{2^{k+1+2t}}{1-2^{1-t}}\log^t(k)\sum_{j=2}^{M}2^{-jt-\frac{k-1}{j}}, \tag{4.16}$$

for any integer $M$ with $3 \leq M \leq 2\sqrt{k-1} - 1$. By Corollary 4.7 we have that

$$jt + \frac{k-1}{j} \geq 2\sqrt{t(k-1)} \quad \forall j, k > 0.$$

Furthermore, we choose $M = \lceil 2\sqrt{\frac{k-1}{t}} + 1 \rceil$. In order to use Proposition 4.18, we need to make sure that $3 \leq M = \lceil 2\sqrt{\frac{k-1}{t}} + 1 \rceil \leq 2\sqrt{k-1} - 1$. $3 \leq \lceil 2\sqrt{\frac{k-1}{t}} + 1 \rceil$ is equivalent to $2 \leq \left\lceil 2\sqrt{\frac{k-1}{t}} \right\rceil$, which holds for $t \leq k-1$ when $k > 1$.
Now we check if $M = \lceil 2\sqrt{\frac{k-1}{t}} + 1 \rceil \leq 2\sqrt{k-1} - 1$. This is equivalent to $\lceil 2\sqrt{\frac{k-1}{t}} \rceil \leq 2\sqrt{k-1} - 2$. For $k \geq 25$, we have $\left\lceil 2\sqrt{\frac{k-1}{t}} \right\rceil \leq \left\lceil 2\sqrt{\frac{k-1}{2}} \right\rceil \leq 2\sqrt{k-1} - 2$.

Our choice $M = \lceil 2\sqrt{\frac{k-1}{t}} + 1 \rceil$ implies that $2\sqrt{\frac{k-1}{t}} + 1 \leq M < 2\sqrt{\frac{k-1}{t}} + 2$, meaning that $M - 1 < 2\sqrt{\frac{k-1}{t}} + 1 < 2\sqrt{\frac{k}{t}} + 1$. Since $t \leq k-1$, we have $1 \leq \sqrt{\frac{k}{t}}$, which yields $M - 1 < 3\sqrt{\frac{k}{t}} < 2^{1.6}\sqrt{\frac{k}{t}}$.
We also see that $1 - M = 1 - (\left\lceil 2\sqrt{\frac{k-1}{t}} \right\rceil + 1) = -\left\lceil 2\sqrt{\frac{k-1}{t}} \right\rceil \leq -2\sqrt{\frac{k-1}{t}}$.
Plugging our chosen value for $M$ and the inequalities established above in (4.16) we get

$$\begin{aligned}
\sum\nolimits'_{n \in M_k} \overline{\alpha}_D(n)^t &\leq 2^{k-2+t(1-M)} \log^t(k) + \frac{2^{k+1+2t}}{1-2^{1-t}} \log^t(k)(M-1)2^{-2\sqrt{t(k-1)}} \\
&< 2^{k-2-2\sqrt{t(k-1)}} \log^t(k) + \frac{2^{k+2.6+2t}}{1-2^{1-t}} \log^t(k)\sqrt{\frac{k}{t}}2^{-2\sqrt{t(k-1)}} \\
&= 2^{k-2-2\sqrt{t(k-1)}} \log^t(k)\left(1 + 2^{4.6}\frac{2^{2t}}{1-2^{1-t}}\sqrt{\frac{k}{t}}\right).
\end{aligned}$$

Let $f(k,t) = \frac{2^{2t}}{1-2^{1-t}}\sqrt{\frac{k}{t}}$. By Lemma 4.20 we know that for a fixed $k_0 \geq 1$, $f(k_0, t)$ is a monotonically increasing function for all $t \geq 2$, and that for a fixed $t_0 \geq 2$, $f(k, t_0)$ is a monotonically increasing function for all $k \geq 1$. Thus, we get with $k \geq 79$ and $t \geq 2$

$$2^{4.6}\frac{2^{2t}}{1-2^{1-t}}\sqrt{\frac{k}{t}} \geq 2^{4.6}\frac{2^4}{1-2^{-1}}\sqrt{\frac{79}{2}} = 4877.38 > 4877.$$

For $x > 4877$ we have $1 + x = x(\frac{1}{x} + 1) < x\frac{4878}{4877}$, which yields

$$\sum\nolimits'_{n \in M_k} \overline{\alpha}_D(n)^t < 2^{k-2-2\sqrt{t(k-1)}} \log^t(k)\frac{4878}{4877}2^{4.6}\frac{2^{2t}}{1-2^{1-t}}\sqrt{\frac{k}{t}}. \tag{4.17}$$

With Corollary 4.6 we uppper bound $2^{-2\sqrt{t(k-1)}}$, which says that for all $t, k \geq 1$ we have $2^{-2\sqrt{t(k-1)}} < 2^{-2\sqrt{tk}} 2^{\sqrt{\frac{t}{k-1}}}$.

For $t = 2$ and $k \geq 88$, and using the fact that $2^{1+\sqrt{\frac{2}{k-1}}}$ is a monotonically decreasing function for all $k \geq 1$, we have

$$\frac{2^{\sqrt{\frac{t}{k-1}}}}{1-2^{1-t}} = \frac{2^{\sqrt{\frac{2}{k-1}}}}{1-2^{1-2}} = 2^{1+\sqrt{\frac{2}{k-1}}} < 2.222.$$

For $3 \leq t \leq k/9$, we have

$$\frac{2^{\sqrt{\frac{t}{k-1}}}}{1-2^{1-t}} \leq \frac{4}{3} 2^{\frac{3}{26}} < 1.7.$$

In any case we have $\frac{2^{\sqrt{\frac{t}{k-1}}}}{1-2^{1-t}} < 2.222$. Putting these estimates in (4.17), we get

$$\begin{aligned} \sum{}'_{n \in M_k} \overline{\alpha}_D(n)^t &< 2^{k-2\sqrt{tk}+2t} \log^t(k) \frac{4878}{4877} 2^{2.6} \frac{2^{\sqrt{\frac{t}{k-1}}}}{1-2^{1-t}} \sqrt{\frac{k}{t}} \\ &< 2^{k-2\sqrt{tk}+2t} \log^t(k) \frac{4878}{4877} 2^{2.6} 2.222 \sqrt{\frac{k}{t}}. \end{aligned}$$

for all $3 \leq t \leq (k-1)/2$, $k \geq 79$ and for $t = 2$, $k \geq 88$.
Now using Proposition 4.9 and (4.2), we get

$$q_{k,t} < \frac{4878}{4877} \frac{2.222}{0.71867} 2^{2.6} \log^t(k) 4^{t-\sqrt{tk}} \frac{k^{3/2}}{\sqrt{t}} < \log^t(k) \frac{k^{3/2}}{\sqrt{t}} 4^{2.12+t-\sqrt{tk}}$$

for $3 \leq t \leq k/9$, $k \geq 21$ and for $t = 2$ $k \geq 88$. □

For the average case error estimate for the Miller-Rabin test we know from [8] that $p_{k,t} < k^{3/2} \frac{2^t}{\sqrt{t}} 4^{2-\sqrt{tk}}$. For $k = 1024$ and $t = 16$ we get:

$$\begin{aligned} q_{1024,16} &< 2.49 \cdot 10^{-54} \text{ using Theorem 4.21} \\ p_{1024,16} &< 7.42 \cdot 10^{-68} \text{ using Theorem 3 in [8].} \end{aligned}$$

### 4.4.4 An estimate for $q_{k,t}$ treating the numbers with large contribution to the estimate differently

We now establish an estimate for $q_{k,t}$ by treating $C_2$, which add the most to our estimate, differently. However, we will see that it is not an improvement over the already established results.

**Theorem 4.22** *For $k \geq 122$ and $t \geq 9$, we have $q_{k,t} < \log^t(k) \Big( (0.35) k 2^{-4t} + \frac{(28.68) 2^{-\frac{k}{2}}}{k} + (3.51) 2^{-t-\frac{k}{3}} k + (3.33) 2^{-2t-\frac{k}{4}} k + (3.30) 2^{-3t-\frac{k}{5}} k$.*

**Proof** By taking $M = 5$ in (4.6), which says $\sum'_{n\in M_k}\overline{\alpha}_D(n)^t \le 2^{t(1-M)}|M_k| + \sum_{m=2}^{M} 2^{(2-m)t}|M_k \cap C_m|$ we have

$$\sum{}'_{n\in M_k}\overline{\alpha}_D(n)^t \le \log^t(k)\Big[2^{-4t} \mid M_k \mid + \mid M_k \cap C_2 \mid + 2^{-t} \mid M_k \cap C_3 \mid + 2^{-2t} \mid M_k \cap C_4 \mid + 2^{-3t} \mid M_k \cap C_5 \mid \Big].$$

We use the bounds $\mid M_k \cap C_m \mid \le c_m 2^{k-\frac{k}{m}}$ established in Lemmas 4.12 and 4.13, which yields in

$$\sum{}'_{n\in M_k}\overline{\alpha}_D(n)^t \le \log^t(k)\Big[2^{k-2-4t} + 20.62\cdot\frac{2^{k/2}}{k^2} + 2.52\cdot 2^{k-\frac{k}{3}} + 2.39\cdot 2^{k-\frac{k}{4}} + 2.37\cdot 2^{k-\frac{k}{5}}.\Big]. \tag{4.18}$$

Now using (4.2) and Proposition 4.9, we get

$$\begin{aligned} q_{k,t} &\le \frac{\sum'_{n\in M_k}\overline{\alpha}_D(n)^t}{\pi(2^k)-\pi(2^{k-1})} \\ &\le \frac{k\log^t(k)}{2^k\cdot 0.71867}\cdot(2^{k-2-4t} + 20.62\cdot\frac{2^{k/2}}{k^2} + 2.52\cdot 2^{k-\frac{k}{3}} \\ &+ 2.39\cdot 2^{k-\frac{k}{4}} + 2.37\cdot 2^{k-\frac{k}{5}}) \\ &\le \log^t(k)\Big((0.35)k2^{-4t} + \frac{(28.68)2^{-\frac{k}{2}}}{k} + (3.51)2^{-t-\frac{k}{3}}k \\ &+ (3.33)2^{-2t-\frac{k}{4}}k + (3.30)2^{-3t-\frac{k}{5}}k\Big). \end{aligned}$$ □

The following result complements Theorems 4.19, 4.21 and 4.22.

**Theorem 4.23** *For integers $k, t$ with $k \ge 122$ and $t \ge k/9$ we have*

$$q_{k,t} < \log^t(k)\Big((0.35)k2^{-4t} + \frac{(28.68)2^{-\frac{k}{2}}}{k} + (3.51)2^{-t-\frac{k}{3}}k + (3.35)2^{-2t-\frac{k}{4}}k\Big).$$

**Proof** We bound the last term of Theorem 4.22 $(3.30)2^{-3t-\frac{k}{5}}$ by $c\cdot 2^{-2t-\frac{k}{4}}$ for some $c \in \mathbb{R}$.

$$(3.30)2^{-3t-\frac{k}{5}} = (3.30)2^{-2t-t-\frac{k}{5}-\frac{k}{20}+\frac{k}{20}} = (3.30)2^{-t+\frac{k}{20}}2^{-\frac{k}{4}-2t}. \tag{4.19}$$

$t \ge k/9$ implies $2^{-t} \le 2^{-\frac{k}{9}}$. With $k \ge 122$ we get

$$\begin{aligned}(3.3)2^{-3t-\frac{k}{5}} &= (3.3)2^{-t+\frac{k}{20}}2^{-\frac{k}{4}-2t} \le (3.3)2^{-\frac{k}{9}+\frac{k}{20}}2^{-\frac{k}{4}-2t} \\ &< 0.0.02^{-\frac{k}{4}-2t}.\end{aligned}$$

Using Theorem 4.22, we get

$$\begin{aligned} q_{k,t} &< \log^t(k)\Big((0.35)k2^{-4t} + \frac{(28.68)2^{-\frac{k}{2}}}{k} + (3.51)2^{-t-\frac{k}{3}}k \\ &\quad + (3.33)2^{-2t-\frac{k}{4}}k + (0.02)2^{-2t-\frac{k}{4}}k\Big) \\ &= \log^t(k)\Big((0.35)k2^{-4t} + \frac{(28.68)2^{-\frac{k}{2}}}{k} + (3.51)2^{-t-\frac{k}{3}}k + (3.35)2^{-2t-\frac{k}{4}}k\Big) \square \end{aligned}$$

Using $k = 1024$ and $t = 115$, we get

$$\begin{aligned} q_{k,t} &< 7.7 \cdot 10^{-38} \\ p_{k,t} &< 2.9 \cdot 10^{-171}. \end{aligned}$$

As $t$ gets larger, the estimate becomes useless. For example for $k = 1024$ and $t = 256$, we already obtain for our estimate from Theorem 4.23 that $q_{k,t} < 3.1 \cdot 10^{12}$.

## 4.5 Bounding $q_{k,l,t}$

We have seen that the estimate of $q_{k,t}$ in Section 4.4 is good when $t$ is small, however is useless for large $t$. Therefore, we are looking for a new bound for $\overline{\alpha}_D(n)$. Let us first state a lemma:

**Lemma 4.24** *Let $n$ be relatively prime to $2D$ and let $\tilde{p}_l$ be the $l$-th prime. If $n$ is not divisible by all of the first $l$ odd primes, then*

$$\varphi_D(n) \leq \left(1 + \frac{1}{\tilde{p}_{l+1}}\right)^{\omega(n)} \cdot n,$$

*which implies*

$$\overline{\alpha}_D(n) \leq \left(1 + \frac{1}{\tilde{p}_{l+1}}\right)^{\omega(n)} \cdot \alpha_D(n).$$

**Proof** For $n_1, n_2 \in \mathbb{N}$ with $\gcd(n_1, n_2) = 1$, we have the relation

$$\varphi_D(n_1, n_2) = \varphi_D(n_1)\varphi_D(n_2).$$

It is thus sufficient to only treat the case $n = p^r$. We have

$$\frac{\varphi_D(p^r)}{p^r} = \frac{p^{r-1}(p - \varepsilon(p))}{p^r} = 1 - \frac{\varepsilon(p)}{p} \leq 1 + \frac{1}{p}.$$

With $p \geq \tilde{p}_{l+1}$, the result follows directly. $\square$

**Example 4.25** *If our odd prime $n$ is not divisible by 3 and relatively prime to $2D$, we get*

$$\overline{\alpha}_D(n) \leq \left(\frac{6}{5}\right)^{\omega(n)} \alpha_D(n). \tag{4.20}$$

**Lemma 4.26** *Let $n \in C_m$. Then*

$$\omega(n) \leq m.$$

**Proof** We know by Lemma 4.3, that

$$\alpha_D(n) \leq 2^{1-\omega(n)} \prod_{i=1}^{\omega(n)} p^{1-r_i} \cdot \frac{(p-\varepsilon(p), n-\varepsilon(n))}{p-\varepsilon(p)} \leq 2^{1-\omega(n)}.$$

Also since $n \in C_m$, we have $2^{-m} < \alpha_D(n)$. Combining the two inequalities yields

$$2^{-m} < \alpha_D(n) \leq 2^{1-\omega(n)},$$

which gives us $\omega(n) - 1 < m$, thus $\omega(n) \leq m$. □

**Definition 4.27** *Let $M_{k,l}$ denote the set of odd $k$-bit integers that are not divisible by the first odd $l$ primes. Let $q_{k,l,t}$ be the probability that a composite integer $n$, which is chosen uniformly at random from $M_{k,l}$ passes $t$ rounds of the strong Lucas test with randomly chosen bases $(P, Q)$.*

With Lemmas 4.24 and 4.26, we are ready for the next theorem.

**Theorem 4.28** *For any integers $k, t, l$ we get*

$$\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t \leq \sum_{m=2}^{\infty} \sum_{n \in M_{k,l} \cap C_m \setminus C_{m-1}} \left(1 + \frac{1}{\tilde{p}_{l+1}}\right)^{mt} 2^{-(m-1)t}.$$

**Proof** By Lemmas 4.24 and 4.26 and the fact that when $m \in C_m \setminus C_{m-1}$ we have that $2^{-m} < \alpha_D(n) \leq 2^{-(m-1)}$, we get

$$\begin{aligned}
\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t &= \sum_{m=2}^{\infty} \sum_{n \in M_{k,l} \cap C_m \setminus C_{m-1}} \overline{\alpha}_D(n)^t \\
&\leq \sum_{m=2}^{\infty} \sum_{n \in M_{k,l} \cap C_m \setminus C_{m-1}} \left(1 + \frac{1}{\tilde{p}_{l+1}}\right)^{\omega(n)t} \alpha_D(n)^t \\
&\leq \sum_{m=2}^{\infty} \sum_{n \in M_{k,l} \cap C_m \setminus C_{m-1}} \left(1 + \frac{1}{\tilde{p}_{l+1}}\right)^{mt} 2^{-(m-1)t}.
\end{aligned}$$ □

**Lemma 4.29** *For any integers $k, t, M, l$ with $3 \leq M \leq 2\sqrt{k-1}-1$, we have*

$$\begin{aligned}\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n)^t \leq\ & 2^{k-2+t} \sum_{m=M+1}^{\infty} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{-mt} \\ & + 2^{k+1+t} \sum_{m=2}^{M} \sum_{j=2}^{m} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{m(1-t)-j-\frac{k-1}{j}}.\end{aligned}$$

**Proof** We know by Theorem 4.28 that

$$\begin{aligned}\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n)^t &\leq \sum_{m=2}^{\infty} \sum_{n\in M_{k,l}\cap C_m\setminus C_{m-1}} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{-(m-1)t} \\ &= \sum_{m=M+1}^{\infty} \sum_{n\in M_{k,l}\cap C_m\setminus C_{m-1}} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{-(m-1)t} \\ &\quad + \sum_{m=2}^{M} \sum_{n\in M_{k,l}\cap C_m\setminus C_{m-1}} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{-(m-1)t} \\ &\leq \sum_{m=M+1}^{\infty} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{-(m-1)t} \mid M_k \mid \\ &\quad + \sum_{m=2}^{M} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{-(m-1)t} \mid M_k \cap C_m \mid \qquad (4.21) \\ &\leq 2^{k-2+t} \sum_{m=M+1}^{\infty} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{-mt} \\ &\quad + 2^{k+1+t} \sum_{m=2}^{M} \sum_{j=2}^{m} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{m(1-t)-j-\frac{k-1}{j}}. \qquad (4.22)\end{aligned}$$

Where the inequality (4.22) follows from Theorem 4.8. □

### 4.5.1 An estimate for $q_{k,l,1}$

Now let us look at the case $t = 1$ in more depth. In order to establish a new estimate, we need the following lemma:

**Lemma 4.30** *For $t = 1$, $3 \leq M \leq 2\sqrt{k-1}-1$, we have*

$$\begin{aligned}\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n) \leq\ & 2^{k-1-M}\Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{M+1} \\ & + 2^{k-2\sqrt{k-1}+1}\Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{M} M(M-1).\end{aligned}$$

**Proof** For $t = 1$ we have from Lemma 4.29 that

$$\begin{aligned}\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n) \leq & 2^{k-1} \sum_{m=M+1}^{\infty} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^m 2^{-m} \\ & + 2^{k+2} \sum_{m=2}^{M} \sum_{j=2}^{m} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^m 2^{-j-\frac{k-1}{j}}.\end{aligned} \tag{4.23}$$

Evaluating the first part of the sum and using that $1+\frac{1}{\tilde{p}_{l+1}} \geq 1$ yields $\sum_{m=M+1}^{\infty} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^m 2^{-m} = \frac{2^{-M}\Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{M+1}}{2-(1+\frac{1}{\tilde{p}_{l+1}})} \leq 2^{-M}\big(1+\frac{1}{\tilde{p}_{l+1}}\big)^{M+1}$.

For the second part of the sum in (4.23) using Corollary 4.7 with $t = 1$, where we have $j+\frac{k-1}{j} \geq 2\sqrt{k-1}$ for all $j$ and $k$, we get that

$$2^{k+2} \sum_{m=2}^{M} \sum_{j=2}^{m} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^m 2^{-j-\frac{k-1}{j}} \leq 2^{k-2\sqrt{k-1}+2} \sum_{j=2}^{M} \sum_{m=j}^{M} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^m.$$

With $m \leq M$, we get

$$\sum_{j=2}^{M} \sum_{m=j}^{M} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^m \leq \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^M \sum_{j=2}^{M} \sum_{m=j}^{M} 1 = \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^M 2^{-1} M(M-1),$$

which concludes our proof. □

**Theorem 4.31** *For $k \geq 2$, we have*

$$q_{k,l,1} < k^2 4^{1.8-\sqrt{k}} \Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^{2\sqrt{k-1}-2}.$$

**Proof** With Lemma 4.30 we get that

$$\begin{aligned}\sum{}'_{n\in M_{k,l}} \overline{\alpha}_D(n)^t \leq & 2^{k-1-M}\big(1+\frac{1}{\tilde{p}_{l+1}}\big)^{M+1} \\ & + 2^{k-2\sqrt{k-1}+1}\Big(1+\frac{1}{\tilde{p}_{l+1}}\Big)^M M(M-1).\end{aligned}$$

Choosing $M = \lfloor 2\sqrt{k-1}-2 \rfloor$, with this we have that $M \geq 2\sqrt{k-1}-3$ and

$M \leq 2\sqrt{k-1} - 2$. Thus we get

$$\begin{aligned}
\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n) &\leq 2^{k-1-M}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{M+1} + 2^{k-2\sqrt{k-1}+1}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{M} M(M-1) \\
&\leq 2^{k-2\sqrt{k-1}+2}\Big(\frac{\tilde{p}_{l+1}+1}{\tilde{p}_{l+1}}\Big)^{M+1} \\
&\quad + 2^{k-2\sqrt{k-1}+1}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{M}(4k - 10\sqrt{k-1} + 2) \\
&\leq 2^{k-2\sqrt{k-1}+2}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{M+1}(1 + 2k - 5\sqrt{k-1} + 1) \\
&\leq 2^{k-2\sqrt{k-1}+2}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{M+1}(2k) \\
&\leq 2^{k-2\sqrt{k-1}+3}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{2\sqrt{k-1}-1} k.
\end{aligned}$$

Combined with inequality (4.2) we have

$$\begin{aligned}
q_{k,l,1} = \frac{\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t}{\pi(2^k) - \pi(2^{k-1})} &\leq \frac{k^2 2^{-2\sqrt{k-1}+3}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{2\sqrt{k-1}-1}}{0.71867} \\
&\leq k^2 4^{-\sqrt{k-1}+1.73}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{2\sqrt{k-1}-1}. \qquad (4.24)
\end{aligned}$$

Let us simplify this expression. Corollary 4.6 says that for all $k \geq 1$ it holds that $2^{-2\sqrt{k-1}} < 2^{-2\sqrt{k}+\frac{1}{\sqrt{k-1}}}$. Thus for $k \geq 53$, we have

$$2^{-2\sqrt{k-1}} < 2^{-2\sqrt{k}+\frac{1}{\sqrt{52}}} < 4^{-\sqrt{k}+0.07}$$

$$q_{k,l,1} < k^2 4^{1.8-\sqrt{k}}\Big(1 + \frac{1}{\tilde{p}_{s+1}}\Big)^{2\sqrt{k-1}-2},$$

which shows that the theorem is true for $k \geq 53$. However,

$$k^2 4^{1.8-\sqrt{k}}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{2\sqrt{k-1}-2} > k^2 4^{1.8-\sqrt{k}} > 1$$

for $k \leq 59$, so the theorem is trivially true for $k \leq 59$. □

Let us look at the bound for $q_{k,l,1}$ in Theorem 4.31 in more detail. When the $(l+1)$-th prime is fairly large, $\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)$ is approximately equal to 1. For example for $k = 1024$, $l = 128$, we get that $\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{2\sqrt{k-1}-1} < 1.09$. When multiplied with the dominant factor $4^{-\sqrt{k}}$, this number is almost negligible. Thus

$$q_{k,l,1} < k^2 4^{1.8-\sqrt{k}}\Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{2\sqrt{k-1}-1} \approx k^2 4^{1.8-\sqrt{k}}.$$

**Example 4.32** *When $k = 1024$, the version 3.0 of OpenSSL checks if our 1024-bit integer is divisible by the first 128 odd primes. The 129-th odd prime is 733. With this we have*

$$q_{1024,128,1} \leq (1024)^2 4^{1.8-\sqrt{1024}} \left(\frac{734}{733}\right)^{2\sqrt{1023}-1} \leq 7.6 \cdot 10^{-13}.$$

*Using Theorem 3 in [8], we would get for $k = 1024$ with the Miller-Rabin test that*

$$p_{1024,1} \leq k^2 4^{2-\sqrt{k}} \leq 9.52 \cdot 10^{-13}.$$

*Thus we already have an improvement over the estimate of the Miller-Rabin test.*

*For $k = 1024$ we also see that this is an improvement over Theorem 4.19, as there we have $q_{1024,1} < 9.6 \cdot 10^{-12}$.*

**Corollary 4.33** *Let $n$ be an odd integer, not divisible by the first 128 odd primes. Then for all $k \geq 2$, we have that $q_{k,128,1} < k^2 4^{1.87727-\sqrt{k}}$.*

**Proof** Using (4.24) we have that

$$q_{k,128,1} \leq k^2 4^{-\sqrt{k-1}+1.73} \left(\frac{734}{733}\right)^{(2\sqrt{k-1}-1)}.$$

We have that

$$\log_4\left(\left(\frac{734}{733}\right)^{(2\sqrt{k-1}-1)}\right) = (2\sqrt{k-1}-1)\log_4\left(\frac{734}{733}\right) \leq (2\sqrt{k-1}-1)0.0009.$$

Thus

$$q_{k,128,1} \leq k^2 4^{1.73-\sqrt{k-1}+0.0009(2\sqrt{k-1}-1)} \leq k^2 4^{1.7291-0.9982\sqrt{k-1}}. \tag{4.25}$$

Using the inequality $\sqrt{k} < 0.9982\sqrt{k-1} + \frac{1}{3\sqrt{k}}$, which holds for all $k \geq 2$, we get for $k \geq 5$ in (4.25)

$$q_{k,128,1} \leq k^2 4^{1.7291-0.9982\sqrt{k-1}} \leq k^2 4^{1.7291-\sqrt{k}+\frac{1}{3\sqrt{k}}} \leq k^2 4^{1.7291-\sqrt{k}+\frac{1}{3\sqrt{5}}} \leq k^2 4^{1.88-\sqrt{k}}.$$

But $k^2 4^{1.88-\sqrt{k}} > 1$ for all $k \leq 60$, so the bound holds trivially for all $k \leq 61$.□

**Example 4.34** *When $k = 2048$, the version 3.0 of OpenSSL checks if our 2048-bit integer is divisible by the first 384 odd primes. The 385-th odd prime is 2659. With this we have*

$$q_{2048,1} \leq (2048)^2 4^{1.8-\sqrt{2048}} \left(\frac{2660}{2659}\right)^{2\sqrt{2047}-1} \leq 2.98 \cdot 10^{-20}.$$

*Using Theorem 3 in [8], we would get for $k = 2048$ with the Miller-Rabin test that*

$$p_{2048,384,1} < 2.99 \cdot 10^{-20}.$$

*We again see that this estimate of the strong Lucas probable prime test is a bit better.*

### 4.5.2 An estimate for $q_{k,l,t}$

Now let let $t \geq 2$. In this section we will establish bounds for $q_{k,l,t}$. Theorem 4.36 is better for smaller $t$, whereas Theorem 4.38 yields a better estimate for larger $t$.

**Corollary 4.35** *Let $\rho = 1 + \frac{1}{\tilde{p}_{l+1}}$. Then*

$$2^t - \rho^t \geq \frac{1}{2}\rho^t.$$

**Proof** Using the fact that $\rho \leq \frac{4}{3} < 2$, we get

$$2^t - \rho^t \geq \rho^t(\frac{2}{\rho} - 1) \geq \rho^t(\frac{2 \cdot 3}{4} - 1) = \rho^t(\frac{3}{2} - 1) = \frac{1}{2}\rho^t.$$

**Theorem 4.36** *For any integers $2 \leq t \leq (k-1)/9$, $k \geq 21, l \in \mathbb{N}$ we have*

$$q_{k,l,t} \leq 4^{1.72-\sqrt{tk}} k^{3/2} 2^t \Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{2\sqrt{kt}+t}.$$

**Proof** By Lemma 4.29, we know that

$$\begin{aligned} \sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t \leq & 2^{k-2+t} \sum_{m=M+1}^{\infty} \Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{-mt} \\ & + 2^{k+1+t} \sum_{j=2}^{M} \sum_{m=j}^{M} \Big(1 + \frac{1}{\tilde{p}_{l+1}}\Big)^{mt} 2^{m(1-t)-j-\frac{k-1}{j}}. \end{aligned} \tag{4.26}$$

for any integer $2 \leq M \leq 2\sqrt{k-1} - 1$. Let us again use the notation $\rho = 1 + \frac{1}{\tilde{p}_{l+1}}$. Let us first look at the left hand side of the sum (4.26). Using Corollary 4.35, which says $2^t - \rho^t \geq \frac{1}{2}\rho^t$, we get that

$$2^{k-2+t} \sum_{m=M+1}^{\infty} \rho^{mt} 2^{-mt} = 2^{k-2+t} \frac{2^{-Mt}\rho^{t(M-1)}}{2^t - \rho^t} \leq 2^{k-2+t} \frac{2^{-Mt}\rho^{t(M-1)}}{2^{-1}\rho^t} \tag{4.27}$$

$$= 2^{k-1-(M-1)t} \rho^{(M-2)t}. \tag{4.28}$$

Now let look at the right hand side of the sum (4.26). Using $\sum_{m=j}^{M} 2^{m(1-t)} < \frac{2^{j(1-t)+t}}{2^t-2}$, and $m \leq M$ we obtain

$$2^{k+1+t} \sum_{j=2}^{M} \sum_{m=j}^{M} \rho^{mt} 2^{m(1-t)-j-\frac{k-1}{j}} \leq \frac{2^{k+1+2t}\rho^{Mt}}{2^t - 2} \sum_{j=2}^{M} 2^{-jt-\frac{k-1}{j}}. \tag{4.29}$$

Now, we shall use Corollary 4.7, which states that $jt + \frac{k-1}{j} \geq 2\sqrt{t(k-1)}$ for all $j, k > 0$. Further, we let $M = \left\lceil 2\sqrt{\frac{k-1}{t}} \right\rceil$. Thus, to have $M \geq 3$, we must restrict $t$ to $t \leq k-1$. Further, for $k \geq 9$, we have

$$M = \left\lceil 2\sqrt{\frac{k-1}{t}} \right\rceil \leq \left\lceil 2\sqrt{\frac{k-1}{2}} \right\rceil \leq 2\sqrt{k-1} - 1.$$

Moreover, we have $M \geq 2\sqrt{\frac{k-1}{t}}$ and $M - 1 < 2\sqrt{\frac{k-1}{t}}$. From (4.26), using (4.27) and (4.29), we get

$$\begin{aligned}
\sum\nolimits'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t &\leq 2^{k-1-(M-1)t} \rho^{(M-2)t} + \frac{2^{k+1+2t-2\sqrt{t(k-1)}}}{2^t - 2} \rho^{Mt} (M-1) \\
&\leq 2^{k-1+t-2\sqrt{t(k-1)}} \rho^{2\sqrt{(k-1)t}-t} \qquad (4.30) \\
&\quad + \frac{2^{k+2+2t-2\sqrt{t(k-1)}}}{2^t - 2} \rho^{2\sqrt{(k-1)t}+t} \sqrt{\frac{k}{t}} \\
&= 2^{k-1+t-2\sqrt{t(k-1)}} \rho^{2\sqrt{(k-1)t}+t} \left( \rho^{-2t} + \frac{2^{3+t}}{2^t - 2} \sqrt{\frac{k}{t}} \right) \\
&< 2^{k-1+t-2\sqrt{t(k-1)}} \rho^{2\sqrt{(k-1)t}+t} \left( 1 + \frac{2^{3+t}}{2^t - 2} \sqrt{\frac{k}{t}} \right). \qquad (4.31)
\end{aligned}$$

The function $\frac{2^t}{2^t-2} \frac{1}{\sqrt{t}}$ is monotonically decreasing for all $t > 1$, thus we have for $t \geq 2$

$$\frac{2^{3+t}}{2^t - 2} \sqrt{\frac{k}{t}} < \frac{2^5}{2} \sqrt{\frac{k}{2}} = \frac{2^4}{\sqrt{2}} \sqrt{k} = 4^{1.75} \sqrt{k}.$$

We have for $k \geq 1$ that $1 + 4^{1.75}\sqrt{k} < \sqrt{k}(1 + 4^{1.75}) = \sqrt{k} 4^{1.812}$. We also use Corollary 4.6, which says that for all $t, k \geq 1$ that $2^{-2\sqrt{t(k-1)}} \leq 2^{-2\sqrt{tk}} 2^{\sqrt{t/(k-1)}}$. For $t \leq (k-1)/9$, we get

$$2^{\sqrt{t/(k-1)}} \leq 2^{\sqrt{1/9}} = 1.25992 < 1.26.$$

Thus, we get from (4.30)

$$\begin{aligned}
\sum\nolimits'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t &\leq 2^{k-1+t} \rho^{2\sqrt{(k-1)t}+t} 4^{1.812-\sqrt{tk}} (1.26) \sqrt{k} \\
&= 2^{k+t} \rho^{2\sqrt{kt}+t} 4^{1.312-\sqrt{tk}} (1.26) \sqrt{k}.
\end{aligned}$$

Using (4.2), we get

$$\begin{aligned}
q_{k,l,t} \leq \frac{\sum\nolimits'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t}{\pi(2^k) - \pi(2^{k-1})} &\leq \frac{2^{k+t} \rho^{2\sqrt{kt}+t} 4^{1.312-\sqrt{tk}} (1.26) \sqrt{k} k}{(0.71867) 2^k} \\
&= 4^{1.72-\sqrt{tk}} k^{3/2} 2^t \Big( 1 + \frac{1}{\tilde{p}_{s+1}} \Big)^{2\sqrt{kt}+t}. \qquad \square
\end{aligned}$$

**Example 4.37** *Let our odd k-bit integer not be divisible by the first* 128 *odd primes. Then Theorem 4.36 says that for* $2 \leq t \leq (k-1)/9$

$$q_{k,128,t} \leq 4^{1.72-\sqrt{tk}} k^{3/2} 2^t \left(\frac{734}{733}\right)^{2\sqrt{kt}+t}.$$

*For* $k = 1024$ *and* $t = 16$*, we get*

$$q_{1024,128,16} \leq 2.92 \cdot 10^{-67}.$$

*Compared with Theorem 3 from [8], which says that* $p_{k,t} < k^{3/2} \frac{2^t}{\sqrt{t}} 4^{2-\sqrt{tk}}$*, we have*

$$p_{1024,16} < 7.42 \cdot 10^{-68}.$$

### 4.5.3 A good estimate when t is large

We are now using a different approach to bound $q_{k,l,t}$, which is a lot more powerful when $t$ is large. In order to achieve such an estimate, we treat the four "worst" categories separately, namely $C_2, C_3, C_4$ and $C_5$. We use Lemmas 4.12 and 4.13, which state that for $k \geq 122$, we have

$$\begin{aligned} \mid M_k \cap C_2 \mid &< 20.62 \frac{2^{k/2}}{k^2} \\ \mid M_k \cap C_3 \mid &\leq (2.52) 2^{k-\frac{k}{3}} \\ \mid M_k \cap C_4 \mid &\leq (2.39) 2^{k-\frac{k}{4}} \\ \mid M_k \cap C_5 \mid &\leq (2.37) 2^{k-\frac{k}{5}}. \end{aligned}$$

**Theorem 4.38** *Let* $k \geq 122$ *and* $\rho = 1 + \frac{1}{\tilde{p}_{s+1}}$*, then*

$$\begin{aligned} q_{k,l,t} \leq 2^{-1.52340-4t} \frac{\rho^{6t}}{2^t - \rho^t} k + \rho^{2t} 2^{4.84257-\frac{k}{2}-t} k^{-1} + \rho^{3t} 2^{1.82-\frac{k}{3}-2t} k \\ + \rho^{4t} 2^{1.74-\frac{k}{4}-3t} k + \rho^{5t} 2^{1.73-\frac{k}{5}-4t} k. \end{aligned}$$

**Proof** Let us choose $M = 5$. With equation (4.21) we get

$$\sum{}'_{n \in M_{k,l}} \overline{\alpha}_D(n)^t \leq 2^{k-2+t} \sum_{m=6}^{\infty} \rho^{mt} 2^{-mt} + \sum_{m=2}^{5} \rho^{mt} 2^{-(m-1)t} \mid M_k \cap C_m \mid . \quad (4.32)$$

Evaluating the first sum yields

$$2^{k-2+t} \sum_{m=6}^{\infty} \rho^{mt} 2^{-mt} = 2^{k-2+t} \frac{\rho^{6t} 2^{-5t}}{2^t - \rho^t} = 2^{k-2-4t} \frac{\rho^{6t}}{2^t - \rho^t}.$$

Thus, we get

$$\begin{aligned}\sum{}'_{n\in M_{k,l}}\overline{\alpha}_D(n)^t \leq & 2^{k-2-4t}\frac{\rho^{6t}}{2^t-\rho^t} + \rho^{2t}2^{-t} \mid M_k \cap C_2 \mid + \rho^{3t}2^{-2t} \mid M_k \cap C_3 \mid \\ & + \rho^{4t}2^{-3t} \mid M_k \cap C_4 \mid + \rho^{5t}2^{-4t} \mid M_k \cap C_5 \mid .\end{aligned} \tag{4.33}$$

Using Lemma 4.12 and 4.13, we get

$$\begin{aligned}\sum{}'_{n\in M_{k,s}}\overline{\alpha}_D(n)^t \leq & 2^{k-2-4t}\frac{\rho^{6t}}{2^t-\rho^t} + \rho^{2t}2^{\frac{k}{2}-t}\frac{20.62}{k^2} + \rho^{3t}2^{k-\frac{k}{3}-2t}(2.52) \\ & + \rho^{4t}2^{k-\frac{k}{4}-3t}(2.39) + \rho^{5t}2^{k-\frac{k}{5}-4t}(3.37).\end{aligned}$$

Using (4.9), we get

$$\begin{aligned}q_{k,l,t} &\leq \frac{\sum{}'_{n\in M_{k,l}}\overline{\alpha}_D(n)^t}{\pi(2^k)-\pi(2^{k-1})} \\ &\leq 2^{-4t}\frac{\rho^{6t}}{2^t-\rho^t}\frac{2^{-2}}{0.71867}k + \rho^{2t}2^{-\frac{k}{2}-t}\frac{20.62}{0.71867}k^{-1} + \rho^{3t}2^{-\frac{k}{3}-2t}\frac{2.52}{0.71867}k \\ &+ \rho^{4t}2^{-\frac{k}{4}-3t}\frac{2.39}{0.71867}k + \rho^{5t}2^{-\frac{k}{5}-4t}\frac{2.37}{0.71867}k \\ &\leq 2^{-1.52340-4t}\frac{\rho^{6t}}{2^t-\rho^t}k + \rho^{2t}2^{4.84257-\frac{k}{2}-t}k^{-1} + \rho^{3t}2^{1.82-\frac{k}{3}-2t}k \\ &+ \rho^{4t}2^{1.74-\frac{k}{4}-3t}k + \rho^{5t}2^{1.73-\frac{k}{5}-4t}k.\end{aligned} \tag{4.34}$$

□

The following Corollary is useful for when $t$ is very large.

**Corollary 4.39** *Let $t \geq k/9$ and $k \geq 122$. Also let $\rho = 1 + \frac{1}{\tilde{p}_{s+1}}$. We then have*

$$q_{k,l,t} \leq 2^{-1.52340-4t}\frac{\rho^{6t}}{2^t-\rho^t}k + \rho^{2t}2^{4.84257-\frac{k}{2}-t}k^{-1} + \rho^{3t}2^{2.14326-\frac{k}{3}-2t}k + \rho^{5t}2^{-3.98-\frac{k}{4}-3t}k.$$

**Proof** For $t \geq k/9$ and $k \geq 122$ the last term of equation (4.34) can be bound through

$$\begin{aligned}\rho^{5t}2^{1.73-\frac{k}{5}-4t} &\leq \rho^{5t}2^{1.73-\frac{k}{4}+\frac{k}{20}-3t-t} \\ &= \rho^{5t}2^{1.73-\frac{k}{4}+\frac{k}{20}-3t-\frac{k}{9}} \\ &\leq \rho^{5t}2^{1.73-\frac{k}{4}-3t-\frac{11k}{180}} \\ &\leq \rho^{5t}2^{1.73-\frac{k}{4}-3t-\frac{11\cdot 122}{180}} \\ &\leq \rho^{5t}2^{-5.72-\frac{k}{4}-3t}.\end{aligned}$$

Thus we get

$$\begin{aligned} q_{k,l,t} &\leq 2^{-1.52340-4t}\frac{\rho^{6t}}{2^t-\rho^t}k + \rho^{2t}2^{4.84257-\frac{k}{2}-t}k^{-1} + \rho^{3t}2^{1.82-\frac{k}{3}-2t}k \\ &+ \rho^{4t}2^{1.74-\frac{k}{4}-3t}k + \rho^{5t}k2^{-5.72-\frac{k}{4}-3t}) \\ &\leq 2^{-1.52340-4t}\frac{\rho^{6t}}{2^t-\rho^t}k + \rho^{2t}2^{4.84257-\frac{k}{2}-t}k^{-1} + \rho^{3t}2^{1.82-\frac{k}{3}-2t}k \\ &+ \rho^{5t}2^{-3.98-\frac{k}{4}-3t}k. \end{aligned}$$

□

### 4.5.4 An estimate excluding the case of twin primes

We are interested in what happens with our average case error probability estimate when we use the fourth variant as discussed in Section 4.3. We make sure that our $n$ is not a product of twin primes, which can for example be achieved using Newton's method. However, we see that this does not make a big difference.

**Corollary 4.40** *Let $n$ be a $k$-bit integer, where $n$ is not of the form $n = m(m+2)$ for an $m \in \mathbb{N}$. Also let $\rho = 1 + \frac{1}{\tilde{p}_{l+1}}$. Then*

$$q_{k,l,t} \leq 2^{-1.52340-4t}\frac{\rho^{6t}}{2^t-\rho^t}k + \rho^{3t}2^{1.82-\frac{k}{3}-2t}k + \rho^{4t}2^{1.74-\frac{k}{4}-3t}k + \rho^{5t}2^{1.73-\frac{k}{5}-4t}k.$$

This follows directly from the proof of Theorem 4.38, where our sum in (4.32) starts at $m = 3$.

Let $k = 1024, t = 115, s = 128, \rho = \frac{734}{733}$. We get

$$\begin{aligned} q_{1024,128,115} &< 5.95 \cdot 10^{-169} \text{ using Theorem 4.38} \\ q_{1024,128,115} &< 1.05 \cdot 10^{-166} \text{ using Theorem 4.36} \\ q_{1024,128,115} &< 5.95 \cdot 10^{-169} \text{ using Corollary 4.40} \\ p_{1024,115} &< 5.06 \cdot 10^{-168} \text{ using Theorem 3 in [8]} \\ p_{1024,115} &< 2.90 \cdot 10^{-171} \text{ using Theorem 6 in [8].} \end{aligned}$$

## 4.6 Outline of the average case proofs of the Miller-Rabin test by Damgård et al.

In Subsection 2.3.1 we have talked about the average case error estimate for the Miller-Rabin test $p_{k,t}$, and stated some bounds established in [8] by Damgård, Landrock and Pomerance: They have shown that $p_{k,1} < k^2 4^{2-\sqrt{k}}$ for all $k \geq 2$, and that $p_{k,t} < k^{3/2}\frac{2^t}{\sqrt{t}}4^{2-\sqrt{tk}}$. We have already mentioned that

the bounds we have obtained for the strong Lucas test have been proven doing a similar analysis. We now give a rough sketch of the proof by Damgård et al. Recall the result from Theorem 2.11, which enabled us to count $S(n)$. Let $\alpha = S(n)/\varphi(n)$ for $n > 1$, $n$ odd and let $B_m$ denote the set of odd composite integers $n$ with $\alpha(n) > 2^{-m}$. They have shown the following theorem:

**Theorem (Theorem 1 in [8])** *If $m, k$ are positive integers with $m + 1 \leq 2\sqrt{k-1}$, then*

$$\frac{\mid B_m \cap M_k \mid}{\mid M_k \mid} < \frac{8}{3}(\pi^2 - 6) \sum_{j=2}^{m} 2^{m-j-(k-1)/j}.$$

Now, let us define $\overline{\alpha}(n) = S(n)/(n-1)$. For the average case error probability we have

$$p_{k,t} = \frac{\sum'_{n \in M_k} \overline{\alpha}(n)^t}{\sum_{n \in M_k} \overline{\alpha}(n)^t} \leq \frac{\sum'_{n \in M_k} \overline{\alpha}(n)^t}{\sum_{p \in M_k} \overline{\alpha}(p)^t} = \frac{\sum'_{n \in M_k} \overline{\alpha}(n)^t}{\pi(2^k) - \pi(2^{k-1})}. \tag{4.35}$$

It is clear that $\varphi(n) \leq n - 1$ for all $n \in N$. Thus, we directly see that $\overline{\alpha}(n) \leq \alpha(n)$. Using this, they have upper bound the final sum in 4.35:

$$\begin{aligned}
\sum'_{n \in M_k} \overline{\alpha}(n)^t &= \sum_{m=3}^{\infty} \sum_{n \in M_k \cap B_m \setminus B_{m-1}} \overline{\alpha}(n)^t \leq \sum_{m=3}^{\infty} \sum_{n \in M_k \cap B_m \setminus B_{m-1}} \alpha(n)^t \\
&\leq \sum_{m=3}^{\infty} 2^{-(m-1)t} \mid M_k \cap B_m \setminus B_{m-1} \mid \\
&\leq 2^{-Mt} \mid M_k \setminus B_M \mid + \sum_{m=3}^{M} 2^{-(m-1)} \mid M_k \cap B_m \mid .
\end{aligned} \tag{4.36}$$

They have bound inequality (4.36) using Theorem 1 in [8] and then they have demonstrated various inequalities in order to establish the average case error bounds. Moreover, in inequality (4.36) they treated $B_3$, which add the most to the estimate for the Miller-Rabin test, differently.

The approach in this thesis broadly follows the same direction. We do numerous adjustments of Theorem 1 from [8], which enabled us to proof Theorem 4.8. As $q_{k,t}$ includes $\overline{\alpha}_D(n)$, we need to find a way to upper bound $\varphi_D(n)$ using $n$. However, we see by Lemma 4.10, that $\varphi_D(n) \leq n$ does not hold in general, therefore we needed to find other ways to do this approximation. Using the estimate by Akbary and Friggstad (see [3]), which says that $\frac{n}{\varphi(n)} \leq (1.07)e^{\gamma} \log(\log(n))$ for $n \geq 2^{78}$, we obtained the bound $\varphi_D(n) < (1.07)n \cdot e^{\gamma} \log(k)$. From there on we presented many mathematical inequalities, which enabled us to prove Theorem 4.19 and Theorem 4.21. However, for large $t$, the theorem is useless, as $\log^t(k)$ increases faster than the rest decreases. Then, we found a way to upper bound $\varphi_D(n)$, namely doing trial division by small primes before performing the strong Lucas test.

This gave us the estimate $\varphi_D(n) \leq \left(1 + \frac{1}{\tilde{p}_{l+1}}\right)^{\omega(n)} n$, where $\tilde{p}_l$ is the $l$-th prime. When $l$ is chosen appropriately, the $\left(1 + \frac{1}{\tilde{p}_{l+1}}\right)^{\omega(n)}$ is close to 1. We again argumented similarly in order to find explicit upper bounds for $q_{k,l,t}$. We then used the same technique to treat the numbers with largest contribution to the estimate for the strong Lucas test differently in the proof.

Chapter 5

# Classifying $C_3$

The numbers that comprise $C_2$ have been characterized in [8]. In Subsections 4.4.4 and 4.5.3 we used this knowledge to get improved estimates for $q_{k,t}$ when $t$ is large, treating $C_2$ differently in the analysis. Now in this section, we classify the members of $C_3$, with the goal of proceeding similarly. However, we will see that the Lucas-Carmichael numbers, see Definition 2.32, belong to this set. Unfortunately, establishing bounds for them is still an open question in number theory, thus we will not be able to proceed further. Once bounds are found, the derivation is straightforward, see inequality (4.33) in the proof of Theorem 4.38.

For the remainder of this chapter, let $D$ be an integer and $n = p_1^{r_1} \dots p_s^{r_s}$ be the prime decomposition of an integer $n$ relatively prime to $2D$. Let $n - \varepsilon(n) = 2^\kappa q$ and $p_i - \varepsilon(p_i) = 2^{k_i} q_i$, with $q, q_i$ odd, ordering the $p_i$'s such that $k_1 \leq \dots \leq k_s$. For integers $m, n, \beta$, we mean by $m^\beta \mid\mid n$ that $m^\beta \mid n$ and $m^{\beta+1} \nmid n$.

Let us first establish all results we need to classify $C_3$. By Lemma 3.3, we know that we have the following inequalities:

$$\frac{SL(D,n)}{\varphi_D(n)} \leq \begin{cases} \frac{1}{2^{s-1}} \prod_{i=1}^{s} \frac{\gcd(q,q_i)}{q_i}, \\ \frac{1}{2^{s-1}} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}}, \\ \frac{1}{2^{s-1+\delta_2+\dots+\delta_s}}, \text{ where } \delta_i = k_i - k_1. \end{cases} \tag{5.1}$$

We also need the following lemmas.

**Lemma 5.1**

$$\frac{SL(D,n))}{\varphi_D(n)} \leq 2^{s+1+\sum_{i=1}^{s}(k_1-k_i)} \prod_{i=1}^{s} \frac{\gcd(q,q_i)}{q_i}.$$

**Proof** From Lemma 4.2 we know that $\left(1+\sum_{j=0}^{k_1-1} 2^{js}\right) \leq 2 \cdot 2^{(k_1-1)s}$. Thus we get

$$\begin{aligned} SL(D,n) =& \Big(\prod_{i=1}^{s} \gcd(q,q_i) - 1\Big) + \sum_{j=0}^{k_1-1} 2^{js} \prod_{i=1}^{s} \gcd(q,q_i) \\ \leq& \Big(1 + \sum_{j=0}^{k_1-1} 2^{js}\Big) \prod_{i=1}^{s} \gcd(q,q_i) \leq 2^{1+(k_1-1)s} \prod_{i=1}^{s} \gcd(q,q_i) \end{aligned}$$

and

$$\varphi_D(n) = \prod_{i=1}^{s} p_i^{r_i-1}(p_i - \varepsilon(p_i)) \geq \prod_{i=1}^{s}(p_i - \varepsilon(p_i)) = \prod_{i=0}^{s} 2^{k_i} q_i.$$

Combining them we get

$$\frac{SL(D,n)}{\varphi_D(n)} \leq 2^{1+(k_1-1)s} \prod_{i=1}^{s} \frac{\gcd(q,q_i)}{2^{k_i} q_i} = 2^{s+1+\sum_{i=1}^{s}(k_1-k_i)} \prod_{i=1}^{s} \frac{\gcd(q,q_i)}{q_i}. \quad (5.2)$$

□

**Lemma 5.2** *Let $n = p_1 p_2$ and $\delta_2 = k_2 - k_1$. Then*

$$2^k q = 2^{2k_1+\delta_2} q_1 q_2 \pm 2^{k_1}(q_1 \pm 2^{\delta_2} q_2).$$

**Proof**

$$\begin{aligned} 2^k q =& p_1 p_2 - \varepsilon(p_1 p_2) \\ =& (2^{k_1} q_1 + \varepsilon(p_1))(2^{k_1+\delta_2} q_2 + \varepsilon(p_2)) - \varepsilon(p_1 p_2) \\ =& 2^{2k_1+\delta_2} q_1 q_2 + 2^{k_1} q_1 \varepsilon(p_2) + 2^{k_1+\delta_2} q_2 \varepsilon(p_1) + \varepsilon(p_1)\varepsilon(p_2) - \varepsilon(p_1 p_2) \\ =& 2^{2k_1+\delta_2} q_1 q_2 + 2^{k_1}(q_1 \varepsilon(p_2) + 2^{\delta_2} q_2 \varepsilon(p_1)) \\ =& 2^{2k_1+\delta_2} q_1 q_2 \pm 2^{k_1}(q_1 \pm 2^{\delta_2} q_2). \end{aligned}$$

□

**Lemma 5.3**

$$\frac{SL(D,n)}{\varphi_D(n)} = \frac{1}{2^{k_1+k_2+\cdots+k_s}} \prod_{i=1}^{s} \frac{1}{p_i^{r_i-1}} \left( \prod_{i=1}^{s} \frac{\gcd(q,q_i) - 1}{q_i} + \frac{2^{sk_1} - 1}{2^s - 1} \prod_{i=1}^{s} \frac{\gcd(q,q_i)}{q_i} \right).$$

**Proof** We have

$$\varphi_D(n) = \prod_{i=1}^{s} \varphi_D(p_i^{r_i}) = \prod_{i=1}^{s} p_i^{r_i-1}(2^{k_i} q_i) = 2^{k_1+k_2+\cdots+k_s} \prod_{i=1}^{s} q_i \prod_{i=1}^{s} p_i^{r_i-1}.$$

Together with

$$\begin{aligned} SL(D,n) =& \Big(\prod_{i=1}^{s} \gcd(q,q_i) - 1\Big) + \sum_{j=0}^{k_1-1} 2^{js} \prod_{i=1}^{s} \gcd(q,q_i) \\ =& \Big(\prod_{i=1}^{s} \gcd(q,q_i) - 1\Big) + \frac{2^{sk_1} - 1}{2^s - 1} \prod_{i=1}^{s} \gcd(q,q_i) \end{aligned}$$

we get the desired result. □

Now we can prove the main theorem of this chapter.

**Theorem 5.4** *The following numbers comprise $C_3$:*

1. $n = 9, 25, 49$.
2. $n = p_1 p_2 = \begin{cases} (2^{k_1} q_1 - 1)(2^{k_1} q_1 + 1), \\ (2^{k_1} q_1 + \varepsilon(p_1))(3 \cdot 2^{k_1} q_1 + \varepsilon(p_2)), \\ (2^{k_1} q_1 + \varepsilon(p_1))(2 \cdot 2^{k_1} q_1 + \varepsilon(p_2)) \text{ with } (q_1, k_1) \neq (1,1), \end{cases}$
   *where all the $q_1$'s are odd and $k_1 \in \mathbb{N}$ and each factor is prime.*
3. $n = 45, 63, 99, 117, 333$.
4. *$n = p_1 p_2 p_3$ is a product of three distinct prime factors, $p_i - \varepsilon(p_i) \mid n - \varepsilon(n)$ and there is some integer $k_1$ such that $2^{k_1} \mid\mid p_i - \varepsilon(p_i)$ for all $i \in \{1,2,3\}$.*

**Proof** 1. Let $s = 1$. Then $n$ is of the form $n = p_1^{r_1}$, where $r_1 \geq 2$. By the second inequality of (5.1), we know that $\alpha_D(n) \leq \frac{1}{p_i^{r_i - 1}}$. Thus if $r_1 \geq 3$, then $\alpha_D(n) \leq \frac{1}{9}$ and $n \notin C_3$. If $r_1 = 2$, then $\alpha_D(n) \leq \frac{1}{11}$ for $p_i > 7$. Thus the only candidates for $n \in C_3$ are $n = 3^2, 5^2, 7^2$.

2. Now let $s = 2$ and $p_1, p_2 \geq 3$. If $p_1 = 3$, then $r_1 \leq 2$ and $r_2 \leq 1$, otherwise by the second inequality of (5.1), we have that $\alpha_D(n) \leq \frac{1}{18}$. If $p_1, p_2 \geq 5$ we again have by the second inequality of (5.1), that $r_i = 1$, because otherwise $\alpha_D(n) \leq \frac{1}{2} \cdot \frac{1}{5} = \frac{1}{10}$. Thus either $n = p_1 p_2$ with $p_1, p_2 \geq 3$ or $n = 3^2 p_2$. The latter case is treated in 3.

   Now let $n = p_1 p_2$ with $p_1 - \varepsilon(p_1) = 2^{k_1} q_1$ and $p_2 - \varepsilon(p_2) = 2^{k_2} q_2$. If $k_2 \geq k_1 + 2$ we have by the third inequality of (5.1), that $\alpha_D(n) \leq \frac{1}{8}$. Thus either $k_2 = k_1$ or $k_2 = k_1 + 1$.

   If $k_1 = k_2$, we have by the first inequality of (5.1) that either both $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$ or $\frac{\gcd(q,q_i)}{q_i} = \frac{1}{3}$ for exactly one $i$ and $\frac{\gcd(q,q_j)}{q_j} = 1$ for the other $j \neq i$, because otherwise $\alpha_D(n) \leq \frac{1}{18}$.

   If $k_2 = k_1 + 1$, it must hold that $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$, otherwise by Lemma 5.2 we have that $\alpha_D(n) \leq \frac{1}{12}$, which implies that $n \notin C_3$.

   For $n = p_1 p_2$, we are left to check the four cases: $k_1 = k_2$ and $q_1 \neq q_2$, $k_1 = k_2$ and $q_1 = q_2$, $k_2 = k_1 + 1$ and $q_1 \neq q_2$ and $k_2 = k_1 + 1$ and $q_1 = q_2$.

   Now let us first assume that $k_1 = k_2$ and $q_1 \neq q_2$. Either both $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$ or one of the fractions is $\frac{1}{3}$ and the other one is 1.

Let $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$, which is equivalent to $q_1, q_2 \mid q$. Thus, $q_1, q_2$ both divide by Lemma 5.2 with $\delta_2 = 0$

$$2^\kappa = 2^{2k_1} q_1 q_2 \pm 2^{k_1}(q_1 \pm q_2).$$

This is only possible if $q_1 = q_2$, which is contradictory to our assumption that $q_1 \neq q_2$.

For the other case, since the primes are ordered with respect to the size of the $k_i$, but here $k_1 = k_2$, the order is irrelevant, we let without loss of generality $q_1 = \gcd(q, q_1)$ and $q_2 = 3\gcd(q, q_2)$. The former is equivalent to $q_1 \mid q$ and the latter is equivalent to $\frac{1}{3}q_2 \mid q$. Thus, $q_1$ and $\frac{1}{3}q_2$ both divide by Lemma 5.2 with $\delta_2 = 0$

$$2^\kappa q = 2^{k_1}(q_1 \pm q_2).$$

It follows that $q_1 \mid q_2$, which implies that there exists some $a \in \mathbb{N}$ such that $q_1 \cdot a = q_2$, and that $\frac{1}{3}q_2 \mid q_2$. This implies that there exists some $b \in \mathbb{N}$ such that $\frac{1}{3}q_2 b = q_1$. Solving the two equations yields in $a = 3$ and $b = 1$, thus $q_2 = 3q_2$. Therefore, $p_1 - \varepsilon(p_1) = 2^{k_1} q_1$ and $p_2 - \varepsilon(p_2) = 2^{k_1} 3q_1$. Thus

$$n = (2^{k_1} q_1 + \varepsilon(p_1))(2^{k_1} 3q_1 + \varepsilon(p_2)) \text{ with } \varepsilon(p_1), \varepsilon(p_2) \in \{\pm 1\}.$$

Now let us check if an $n$ of such a form is actually in $C_3$. By Lemma 5.3 we have

$$\alpha_D(n) = \frac{1}{4^{k_1}}\Big(\frac{q_1 - 1}{q_1}\frac{\frac{1}{3}q_2 - 1}{q_2} + \frac{4^{k_1} - 1}{9}\Big) = \frac{1}{4^{k_1}}\Big(\Big(\frac{q_1 - 1}{q_1}\Big)^2 \cdot \frac{1}{3} + \frac{4^{k_1} - 1}{9}\Big).$$

We consider two cases: $q_1 = 1$ and $q_1 \neq 1$. Let $q_1 = 1$. Since $\frac{4^{k_1}-1}{4^{k_1}} < 1$, we get $\alpha_D(n) = \frac{4^{k_1}-1}{4^{k_1}\cdot 9} < \frac{1}{8}$, so $n \notin C_3$. Now let $q_1 \neq 1$. Thus, $\alpha_D(n) = \frac{1}{4^{k_1}}\Big(\Big(\frac{q_1-1}{q_1}\Big)^2 \cdot \frac{1}{3} + \frac{4^{k_1}-1}{9}\Big) \geq \frac{1}{4^{k_1}\cdot 3}\Big(\frac{1}{4} + \frac{4^{k_1}-1}{3}\Big) = \frac{1}{3\cdot 4^{k_1}}\frac{4^{k_1+1}-1}{12} \geq \frac{5}{48} > \frac{1}{8}$, where the inequalities follow from the fact that boht $\frac{q_1-1}{q_1}$ and $\frac{4^{k_1+1}-1}{4^{k_1}}$ are monotonically increasing functions in $q_1$ and $k_1$ respectively. Thus $n \in C_3$.

Now let $k_2 = k_1$ and $q_1 = q_1$. Then $p_1 - \varepsilon(p_1) = 2^{k_1} q_1$ and $p_2 - \varepsilon(p_2) = 2^{k_1} q_1$. In order for $p_1$ and $p_2$ to be distinct primes, we must have that $\varepsilon(p_1) \neq \varepsilon(p_2)$. Without loss of generality we assume that $\varepsilon(p_1) = 1$ and thus $\varepsilon(p_2) = -1$. Therefore, $n = (2^{k_1} q_1 - 1)(2^{k_1} q_1 + 1)$. By Lemma 3.11, we know that $\frac{SL(D,n)}{\varphi_D(n)} > \frac{1}{3}$ for all odd $q_1 \neq 1$. When $q_1 = 1$, then $\frac{SL(D,n)}{\varphi_D(n)} = \frac{1}{3} - \frac{1}{3\cdot 4^{k_1}}$. Since $-\frac{1}{3\cdot 4^{k_1}}$ is monotonically increasing in $k_1$, we have that $\alpha_D(n) = \frac{1}{3} - \frac{1}{3\cdot 4^{k_1}} \geq \frac{1}{3} - \frac{1}{3\cdot 4^1} = \frac{1}{4} > \frac{1}{8}$. Thus, $n \in C_3$.

Now let $k_2 = k_1 + 1$, which means that $\delta_2 = 1$, and let $q_1 \neq q_2$. We know that $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$, which is equivalent to $q_1 \mid q$ and $q_2 \mid q$. Therefore, $q_1, q_2$ both divide by Lemma 5.2 with $\delta_2 = 1$

$$2^k q = 2^{2k_1+1} q_1 q_2 \pm 2^{k_1}(q_1 \pm 2q_2).$$

Thus it must hold that $q_1 \mid 2q_2$ and $q_2 \mid q_1$. Since $q_1$ is odd, we must have that $q_1 \mid q_2$, which is only possible when $q_1 = q_2$, which is a contradiction to our assumption that $q_1 \neq q_2$.

Now let us consider the remaining possible case, namely when $k_2 = k_1 + 1$ and $q_1 = q_2$. We again have that $q_1 \mid q$ and $q_2 \mid q$. Therefore, $p_1 - \varepsilon(p_1) = 2^{k_1} q_1$ and $p_2 - \varepsilon(p_2) = 2^{k_1+1} q_1 = 2(2^{k_1} q_1) = 2(p_1 - \varepsilon(p_1))$. Therefore,

$$n = p_1 p_2 = (2^{k_1} q_1 + \varepsilon(p_1))(2 \cdot 2^{k_1} q_1 + \varepsilon(p_2)) \text{ with } \varepsilon(p_1), \varepsilon(p_2) \in \{\pm 1\}.$$

Let us check if such an $n$ is in $C_3$. By Lemma 5.3, we have that

$$\alpha_D(n) = \Big(\frac{q_1 - 1}{q_1}\Big)^2 \cdot \frac{1}{2 \cdot 4^{k_1}} + \frac{4^{k_1} - 1}{6 \cdot 4^{k_1}}.$$

We consider two cases: $q_1 = 1$ and $q_1 \neq 1$. With $q_1 = 1$, we obtain $\alpha_D(n) = \frac{4^{k_1}-1}{6 \cdot 4^{k_1}}$. This is $> \frac{1}{8}$ if and only if $k_1 > 1$. For $k_1 = 1$, we obtain $\alpha_D(n) = \frac{1}{8}$, the only possibility is $n = (2 + \varepsilon_1)(4 + \varepsilon_2) = 3 \cdot 5$. With $q_1 \neq 1$ and the fact that $(q_1 - 1)/q_1$ is monotonically increasing, we obtain

$$\begin{aligned}
\alpha_D(n) &= \Big(\frac{q_1 - 1}{q_1}\Big)^2 \cdot \frac{1}{2 \cdot 4^{k_1}} + \frac{4^{k_1} - 1}{6 \cdot 4^{k_1}} \\
&\geq \frac{4}{9} \cdot \frac{1}{2 \cdot 4^{k_1}} + \frac{4^{k_1} - 1}{4^{k_1} \cdot 6} = \frac{1}{3}\Big(\frac{2}{3 \cdot 4^{k_1}} + \frac{4^{k_1} - 1}{4^{k_1} \cdot 2}\Big) \\
&= \frac{1}{3}\Big(\frac{3 \cdot 4^{k_1} + 1}{4^{k_1} \cdot 6}\Big) = \frac{1}{6} + \frac{1}{18 \cdot 4^{k_1}} > \frac{1}{8}.
\end{aligned}$$

3. Let $s = 2$ and this time $p_1 = 3$ and $r_1 = 2$, meaning $n = 3^2 p_2$. Since $3 - \varepsilon(3) = 2^{k_1} q_1$, but $\varepsilon(3) = \pm 1$, we have that $3 - \varepsilon(3) \in \{2, 4\}$, which implies that $q_1 = 1$ and $k_1 \in \{1, 2\}$.
   By the third inequality of (5.1), we have for $k_2 \geq k_1 + 2$ that $n \notin C_3$, thus either $k_1 = k_2$ or $k_2 = k_1 + 1$ŋewline Now let $k_1 = k_2$. Again it must hold that either $\frac{\gcd(q,q_1)}{q_1} = \frac{\gcd(q,q_2)}{q_2} = 1$ or $\frac{\gcd(q,q_1)}{q_1} = 1$ and $\frac{\gcd(q,q_2)}{q_2} = 3$, since $q_1 = 1$.

We have

$$\begin{aligned}
2^\kappa q =& n - \varepsilon(n) = 3^2 p_2 - \varepsilon(3^2 p_2) \\
=& (2^{k_1} + \varepsilon(3))^2 (2^{k_1+\delta_2} q_2 + \varepsilon(p_2)) - \varepsilon(p_2) \\
=& (2^{2k_1} + 2^{k_1+1}\varepsilon(3) + 1)(2^{k_1+\delta_2} q_2 + \varepsilon(p_2)) - \varepsilon(p_2) \\
=& (2^{3k_1+\delta_2} q_2 + 2^{2k_1+1+\delta_2} q_2 \varepsilon(3) + 2^{k_1+\delta_2} q_2 + \varepsilon(p_2)(2^{2k_1} + 2^{k_1+1}\varepsilon(3)) \\
& + \varepsilon(p_2) - \varepsilon(p_2) \\
=& q_2((2^{3k_1+\delta_2} + 2^{2k_1+1+\delta_2}\varepsilon(3) + 2^{k_1+\delta_2}) \pm (2^{2k_1} + \varepsilon(3)2^{k_1+1}). \qquad (5.3)
\end{aligned}$$

Now let us look at the case where $\frac{\gcd(q,q_2)}{q_2} = 1$, meaning $q_2 \mid q$. With this $q_2 \mid 2^\kappa q$, thus it must also divide (5.3), which implies that $q_2 \mid 2^{2k_1} \pm 2^{k_1+1}$. Since $k_1 \in \{1,2\}$, it follows that for $k_1 = 1$ either $q_2 \mid 8$ if $\varepsilon(3) = 1$ or $q_2 \mid 0$ if $\varepsilon(3) = -1$, and for $k_1 = 2$ either $q_2 \mid 24$ if $\varepsilon(3) = 1$ or $q_2 \mid 8$ if $\varepsilon(3) = -1$. Since $q_2$ must be odd, the only possibilities are when $(k_1, q_2, \varepsilon(3)) = (1,1,1), (k_1, q_2, \varepsilon(3)) = (2,1,\pm 1)$ or $(k_1, q_2, \varepsilon(3)) = (2,3,1)$. However, with $k_1 = 1$, it must hold that $\varepsilon(3) = 11$, otherwise $2^{k_1} + \varepsilon(3) \neq 3$ and for $k_1 = 2$, it must hold that $\varepsilon(3) = -1$, otherwise $2^{k_1} + \varepsilon(3) \neq 3$. Thus we are only left with $(k_1, q_2, \varepsilon(3)) = (1,1,1)$ and $(k_1, q_2, \varepsilon(3)) = (2,1,-1)$. This analysis holds for both $k_2 = k_1$ and $k_2 = k_1 + 1$. Therefore, we get

$$p_2 = 2^{k_2} q_2 + \varepsilon(p_2) =$$
$$\begin{cases}
2^{k_1} q_2 \pm 1 = 2^1 \cdot 1 \pm 1 = 1, 3, & \text{if } k_1 = k_2 = 1, q_2 = 1, \varepsilon(3) = 1 \\
2^{k_1} q_2 \pm 1 = 2^2 \cdot 1 \pm 1 = 3, 5, & \text{if } k_1 = k_2 = 2, q_2 = 1, \varepsilon(3) = -1 \\
2^{k_1+1} q_2 \pm 1 = 2^2 \cdot 1 \pm 1 = 3, 5 & \text{if } k_2 = k_1 + 1, k_1 = 1, q_2 = 1, \varepsilon(3) = 1 \\
2^{k_1+1} q_2 \pm 1 = 2^3 \cdot 1 \pm 1 = 7, 9 & \text{if } k_2 = k_1 + 1, k_1 = 2, q_2 = 1, \varepsilon(3) = -1.
\end{cases}$$

Since $p_2$ is a prime different from 3, we discard all other cases and are left with $p_2 \in \{5,7\}$.
Now let us look at the case where $\frac{\gcd(q,q_2)}{q_2} = \frac{1}{3}$, meaning $\frac{1}{3}q_2 \mid q$. Here it must hold that $k_1 = k_2$. By the same reasoning as above we have $\frac{1}{3}q_2 \mid 2^{2k_1} \pm 2^{k_1+1}$, which implies $q_2 \mid 3(2^{2k_1} \pm 2^{k_1+1})$. For $k_1 = 1$, we have either $q_2 \mid 24$ if $\varepsilon(3) = 1$ or $q_2 \mid 0$ if $\varepsilon(3) = -1$, for $k_1 = 2$, we have either $q_2 \mid 72$ if $\varepsilon(3) = 1$ or $q_2 \mid 24$ if $\varepsilon(3) = -1$. Again since $q_2$, must be odd, the only possible combinations are for $(k_1, q_2, \varepsilon(3)) = (1,1,1), (k_1, q_2, \varepsilon(3)) = (1,3,1)$ or $(k_1, q_2, \varepsilon(3)) \in$

$\{(2,1,1),(2,1,\pm 1),(2,3,\pm 1),(2,9,1)\}$. Thus we get

$$p_2 = 2^{k_2}q_2 + \varepsilon(p_2) =$$
$$\begin{cases} 2^{k_1}q_2 \pm 1 = 2\cdot 1 \pm 1 = 1,3 & \text{if } k_1 = 1, q_2 = 1 \text{ and } \varepsilon(3) = 1 \\ 2^{k_1}q_2 \pm 1 = 2\cdot 3 \pm 1 = 5,7, & \text{if } k_1 = 1, q_2 = 3 \text{ and } \varepsilon(3) = 1 \\ 2^{k_1}q_2 \pm 1 = 2^2\cdot 1 \pm 1 = 3,5, & \text{if } k_1 = 2, q_2 = 1 \text{ and } \varepsilon(3) = \pm 1 \\ 2^{k_1}q_2 \pm 1 = 2^2\cdot 3 \pm 1 = 11,13, & \text{if } k_1 = 2, q_2 = 3 \text{ and } \varepsilon(3) = \pm 1 \\ 2^{k_1}q_2 \pm 1 = 2^2\cdot 9 \pm 1 = 35,37, & \text{if } k_1 = 2, q_2 = 9 \text{ and } \varepsilon(3) = 1. \end{cases}$$

Again we discard the cases where $p_2 = 1$, composite or divisible by 3 and are left with $p_2 \in \{5,7,11,13,37\}$.
We see that for $n = 3^2 p_2$ with $p_2 \geq 5$ prime and $n \in C_3$, $n$ must be of the form: $n \in \{3^2\cdot 5, 3^2\cdot 7, 3^2\cdot 11, 3^2\cdot 13, 3^2\cdot 37\} = \{45,63,99,117,333\}$.

4. Now let $s = 3$ with $n = p_1^{r_1}p_2^{r_2}p_3^{r_3}$. By the second inequality of (5.1), it must hold that $r_i = 1$ for all $i \in \{1,2,3\}$, otherwise we would have $\alpha_D(n) \leq \frac{1}{12}$. Therefore, $n = p_1p_2p_3$ with $p_i \neq p_j \ \forall i \neq j$. By the first inequality of (5.1), we have that $\frac{\gcd(q,q_i)}{q_i} = 1$ for all $i \in \{1,2,3\}$, otherwise $\alpha_D(n) \leq \frac{1}{12}$, in which case we would have that $n \notin C_3$. Thus we must have that $q_i \mid q \ \forall i \in \{1,2,3\}$. By the third inequality of (5.1), we must have that $k_1 = k_2 = k_3$, as else $\alpha_D(n) \leq \frac{1}{8}$.
Therefore, we have $k_1 = k_2 = k_3$ with $q_i \mid q$ for all $i \in \{1,2,3\}$. It is clear that $q \mid 2^\kappa q$, thus $q_i \mid 2^\kappa q$. But we know by Lemma 4.1 that $2^{k_i} \mid 2^\kappa$ for every $i$, this implies that $2^{k_i}q_i \mid 2^\kappa q$, which is the same as saying that $p_i - \varepsilon(p_i) \mid n - \varepsilon(n)$.

Let us check if such an $n$ is indeed in $C_3$. Using Lemma 5.3 and the fact that $k_1 = k_2 = k_3$, $q_i \mid q$ and $r_i = 1$ for $i = 1,2,3$ we get

$$\begin{aligned} \alpha_D(n) &= \frac{1}{2^{3k_1}}\left(\prod_{i=1}^{3}\frac{q_i - 1}{q_i} + \frac{2^{3k_1}-1}{7}\right) \\ &= \frac{1}{2^{3k_1}}\prod_{i=1}^{s}\frac{q_i-1}{q_i} + \frac{1}{7}\cdot\frac{2^{3k_1}-1}{2^{3k_1}}. \end{aligned}$$

Since $\frac{2^{3k_1}-1}{3k_1}$ is monotonically increasing in $k_1$, we get $\frac{2^{3k_1}-1}{3k_1} \geq \frac{2^3-1}{2^3} = \frac{7}{8}$. Thus

$$\begin{aligned} \alpha_D(n) &= \frac{1}{2^{3k_1}}\prod_{i=1}^{s}\frac{q_i-1}{q_i} + \frac{1}{7}\cdot\frac{2^{3k_1}-1}{2^{3k_1}} \\ &\geq \frac{1}{2^{3k_1}}\prod_{i=1}^{s}\frac{q_i-1}{q_i} + \frac{1}{8} > \frac{1}{8}. \end{aligned}$$

With this we indeed have that $n \in C_3$.

5. Now let $s \geq 4$. By the second inequality of (5.1), we immediately have that $\alpha_D(n) \leq \frac{1}{8}$, thus $n \notin C_3$. □

The numbers of the fourth form in Theorem 5.4 are by Theorem 2.33 the Lucas-Carmichael numbers with three prime factors, that have the additional property that there exists some $k_1 \in \mathbb{N}$ such that $2^{k_1} \mid\mid p_i - \varepsilon(p_i)$ for all $i = 1, 2, 3$.

Chapter 6

# Further investigations on the strong Lucas probable prime test

We know by Theorem 3.12 that $SL(D,n) \leq \frac{4n}{15}$ for every odd composite integer $n$ which is not a product of twin-primes. Therefore it is tempting to directly conclude that $q_{k,t} \leq (\frac{4}{15})^t$. For the same reason as in Subsection 2.3.1, we cannot proceed in this way. However, if we first check that $n \neq p(p+2)$ for some $p \in \mathbb{N}$ and $n$ is not divisible by the first 128 odd primes, then for $k \geq 67$, we indeed have $q_{k,t} \leq \left(\frac{4}{15}\right)^t$. If we relax the assumption of not being divisible by small primes, we can show that $q_{k,t} \leq \left(\frac{4}{15}\right)^t$ for each $k \geq 111$.

Let $X$ represent the event that an integer $n$, which is not a product of twin-primes, is composite, let $D_i$ denote the event that an integer chosen at random from $M_k$ passes the $i$-th strong Lucas test, and let $Z_t$ denote the event that it passes $t$ consecutive rounds of the test, i.e. $Z_t = D_1 \cap D_2 \cap \cdots \cap D_t$. Thereom 3.12 states that $\mathbb{P}[Z_t \mid X] \leq \left(\frac{4}{15}\right)^t$, and what is relevant is $q_{k,t} = \mathbb{P}[X \mid Z_t]$. The next lemma and its proof is based on [6] adapted for the strong Lucas test.

**Lemma 6.1** *Let $n$ be an odd integer, that is not a product of twin primes and let $j, t \in \mathbb{Z}^+$ with $1 \leq j \leq t-1$. Then*

$$q_{k,t} \leq \left(\frac{4}{15}\right)^{t-j} \frac{q_{k,j}}{1-q_{k,j}}.$$

**Proof** We let $X'$ denote set of composites and $\overline{\alpha}_D(n) = SL(D,n)/n$. Recall that for odd composite $n$ we will have have $\overline{\alpha}_D(n) \leq 4/15$. We note that $\mathbb{P}[X \cap Z_i] = 2^{-(k-2)} \sum_{n \in X' \cap M_k} \overline{\alpha}_D(n)^i$.

For $1 \leq j \leq t-1$, we have

$$q_{k,t} = \mathbb{P}[X \mid Z_t] = \frac{\mathbb{P}[X \cap Z_t]}{\mathbb{P}[Z_t]}$$
$$= \frac{\mathbb{P}[X \cap E_t]}{X \cap E_{t-1}} \frac{\mathbb{P}[X \cap E_{t-1}]}{X \cap E_{t-2}} \cdots \frac{\mathbb{P}[X \cap E_{j+1}]}{X \cap E_l} \frac{\mathbb{P}[X \cap E_j]}{X \cap E_t}.$$

Now

$$\frac{\mathbb{P}[X \cap E_i]}{X \cap E_{i-1}} = \frac{\sum_{n \in X' \cap M_k} \overline{\alpha}_D(n)^i}{\sum_{n \in X' \cap M_k} \overline{\alpha}_D(n)^{i-1}} \leq \frac{\sum_{n \in X' \cap M_k} \frac{4}{15} \overline{\alpha}_D(n)^{i-1}}{\sum_{n \in X' \cap M_k} \overline{\alpha}_D(n)^{i-1}} = \frac{4}{15}.$$

Let $X^{\mathsf{c}}$ denote the complement of the event $X$, which expresses the event that a number is prime. Since a prime in $M_k$ always passes each strong Lucas test, we see that $\mathbb{P}[X^{\mathsf{c}} \cap Z_t] = \mathbb{P}[X^{\mathsf{c}}] = \mathbb{P}[X^{\mathsf{c}} \cap Z_j]$. Therefore,

$$q_{k,t} \leq \left(\frac{4}{15}\right)^{t-j} \frac{\mathbb{P}[X \cap Z_j]}{\mathbb{P}[Z_j]} \frac{\mathbb{P}[Z_j]}{\mathbb{P}[Z_t]} = \left(\frac{4}{15}\right)^{t-l} q_{k,j} \frac{\mathbb{P}[Z_j]}{\mathbb{P}[Z_t]}.$$

Thus,

$$\frac{\mathbb{P}[Z_j]}{\mathbb{P}[Z_t]} \leq \frac{\mathbb{P}[Z_j]}{\mathbb{P}[X^{\mathsf{c}} \cap Z_t]} = \frac{\mathbb{P}[Z_j]}{\mathbb{P}[X^{\mathsf{c}} \cap Z_j]} = \frac{1}{\mathbb{P}[X^{\mathsf{c}} \mid Z_j]} = \frac{1}{1 - q_{k,t}},$$

which completes the proof of the lemma. □

**Theorem 6.2** *Let the odd integer tested for primality not be a product of twin primes and not be divisible by the first 128 odd primes. Then for all $t \geq 1$ and $k \geq 67$ we have*

$$q_{k,t} \leq \left(\frac{4}{15}\right)^t.$$

**Proof** Taking $j = 1$ in Lemma 6.1 we get that $q_{k,t} \leq \left(\frac{4}{15}\right)^{t-1} \frac{q_{k,j}}{1-q_{k,j}}$. So to show that $q_{k,t} \leq \left(\frac{4}{15}\right)^t$ it thus suffices to show that $q_{k,j} \leq 4/19$. By Corollary 4.33 we have $q_{k,1} < k^2 4^{1.87727-\sqrt{k}}$. With this we can easily calculate that $q_{k,j} \leq \frac{4}{19}$ for each $k \geq 91$. By taking $j = 2$ in Lemma 6.1 we get that $q_{k,t} \leq \left(\frac{4}{15}\right)^{t-2} \frac{q_{k,2}}{1-q_{k,2}}$ for $t \geq 3$. With this it suffices to show that $q_{k,1} \leq 4/15$ and $q_{k,2} \leq 16/225$. Using Theorem 4.36 with $t = 2$ we get for all $k \geq 21$ that $q_{k,2} \leq 4^{1.68617-\sqrt{2k}} k^{3/2} 4 \cdot \left(\frac{734}{733}\right)^{2\sqrt{k2}+2}$, so that $q_{k,2} \leq 16/225$ for each $k \geq 37$. By Lemmas 4.30, 4.9 and inequality (4.2) with $\tilde{p}_{l+1} = 733$ and choosing $M = 10$, we get that

$$q_{k,1} \leq \frac{\sum'_{n \in M_k} \overline{\alpha}_D(n)}{\pi(2^k) - \pi(2^{k-1})} \leq \frac{2^{-1-M}(\frac{734}{733})^{M+1} + 2^{-2\sqrt{k-1}+1} + (\frac{734}{733})^M M(M-1)}{0.71867}.$$

With this we have $q_{k,1} \leq 4/15$ for all $k \in \{67, \ldots, 386\}$. So we have shown that $q_{k,t} \leq \left(\frac{4}{15}\right)^t$ for all $k \geq 67$, which completes the proof. □

**Theorem 6.3** *Let the integer tested for primality not be a product of twin primes Then* $q_{k,t} \leq \left(\frac{4}{15}\right)^t$ *for each* $t \geq 1$ *and* $k \geq 111$.

**Proof** The proof is identical as the proof from Theorem 6.2. By Theorem 4.19, we have for $k \geq 2$ that $q_{k,1} < \log(k)k^2 4^{2.3-\sqrt{k}}$, so $q_{k,1} \leq 4/19$ for each $k \geq 136$, thus by Lemma 6.1 we have that $q_{k,t} \leq \left(\frac{4}{15}\right)^t$. Now we show that $q_{k,1} \leq 4/15$ and $q_{k,2} \leq 16/225$ for $k \geq 111$, which again by Lemma 6.1 is sufficient to conclude that $q_{k,t} \leq \left(\frac{4}{15}\right)^t$. By inequalty (4.11), we have that $\sum'_{n \in M_k} \overline{\alpha}_D(n) \leq \log(k)\left(2^{k-1-M} + 2^{k+3-2\sqrt{k-1}} \sum_{j=2}^{M}(M+1-j)\right)$, combined with inequality (4.2), we have

$$q_{k,1} \leq \frac{\sum'_{n \in M_k} \overline{\alpha}_D(n)}{\pi(2^k) - \pi(2^{k-1})} \leq \frac{\log(k) \cdot k\left(2^{-1-M} + 2^{3-2\sqrt{k-1}} \sum_{j=2}^{M}(M+1-j)\right)}{0.71867}.$$

Choosing $M = 13$, we get that $q_{k,1} \leq 4/15$ for $k \in \{111, \ldots 504\}$. Now by Theorem 4.21 we have for $k \geq 88$, $t = 2$ that $q_{k,2} < \frac{k^{3/2}}{\sqrt{t}} \log^t(k) 4^{4.615-\sqrt{2k}}$. So we get that $q_{k,2} < 16/225$ for $k \geq 88$, which concludes the proof. □

Chapter 7

# Conclusion

In this Master's thesis, we established the framework needed to find average case error bounds for the strong Lucas test: We considered a procedure, which chooses randomly $k$-bit integers from the uniform distribution, subjects each number to $t$ iterations of the strong Lucas test and outputs the first number that passes all $t$ tests. Let $q_{k,t}$ be the probability, that this procedure outputs a composite integer. The bounds we obtained are $q_{k,1} \leq \log(k)k^2 4^{2.3-\sqrt{k}}$ for $k \geq 2$ and $q_{k,t} < \log^t(k)\frac{k^{3/2}}{\sqrt{t}}4^{2.12+t-\sqrt{tk}}$ for $k \geq 21$ and $t \geq 2$. If the integer is divisible by a small prime it is computationally less expensive to rule out the candidate by trial division than by using the strong Lucas test. We saw that the bounds of the procedure improved after we imposed the additional requirement to check for divisibility by the first odd $l$-primes in the step before running the strong Lucas test. Let $q_{k,l,t}$ be the probability that this updated procedure returns a composite number. Let $\tilde{p}_l$ denote the $l$-th odd prime. We showed that $q_{k,l,1} < k^2 4^{1.8-\sqrt{k}}\left(1+\frac{1}{\tilde{p}_{l+1}}\right)^{2\sqrt{k-1}-2}$ for all $l \in N$ and $k \geq 1$ and $q_{k,l,t} \leq 4^{1.68617-\sqrt{tk}}k^{3/2}2^t\left(1+\frac{1}{\tilde{p}_{l+1}}\right)^{2\sqrt{kt}+t}$ for all $k \geq 21$ and $t \geq 2$.

The bounds using the second case are comparable to the bounds in [8] of the Miller-Rabin test. Next, we did an error analysis treating the numbers that add the most to our probability estimate differently, which resulted in good bounds for large $t$.

We then classified the integers with the second largest contribution to our estimate, hoping to do a similar analysis as we did for the numbers with the largest contribution. Unfortunately, Lucas-Carmichael numbers belong to this set. This is problematic as there is still the open question of bounding the number of Lucas-Carmichael numbers less than a given integer $x$. Thus, were were not able to proceed further.

In the last chapter, we showed that for odd composite integers $n$, which are not a product of twin primes, the bound $q_{k,t} \leq (4/15)^t$ holds for $k \geq 111$. This result is remarkable because it does not follow directly from $SL(D, n) \leq 4/15$, as it does not take the distribution of primes into account. Of course as stated above, we were able to show that $q_{k,t}$ is, in fact, smaller than $(4/15)^t$ for all sufficiently large $k$.

During the scope of this work, average case error bounds for the strong Lucas test were found. Yet, many open questions that look promising for future projects remain. For example, future works could be to bound the set of odd $k$-bit integers, which are not divisible by the first odd $s$ primes. This could result in improved estimates for the strong Lucas test. Once there exist bounds for the number of Lucas-Carmichael numbers, one could tighten the established bounds for large $t$. Moreover, one could try to obtain bounds for the average case error probability for the "normal" Lucas test and investigate bounds using incremental search for the (strong) Lucas test. Furthermore, one could analyse if it is possible to get improved estimates for the Miller-Rabin test using the modified procedure that includes division by small primes. The most interesting future work however, is to get average case error bounds for the Baillie-PSW test.

# Bibliography


[1] P. van Oorschot A. Menezes, S. Vanstone. *Handbook of Applied Cryptography.* CRC Press, 1997.

[2] M. Agrawal and S. Biswas. *Primality and identity testing via chinese remaindering*. Journal of the ACM, 50:429–443, 1999.

[3] A. Akbary and Z. Friggstad. *Explicit upper bounds for* $\prod_{p \leq p_{\omega(n)}} \frac{p}{p-1}$. Mathematics Subject Classification: Primary 11Y70, 11A25., 2000.

[4] Arnault. *The Rabin-Monier Theorem for Lucas Pseudoprimes*. Mathematics of Computation, Vol 66, Number 218, 1997.

[5] F. Arnault. *Sur quelques tests probabilistes de primalité*. Doctoral thesis, l'Université de POITIERS, 1992.

[6] R. Burthe. *Further investigations with the strong probable prime test*. Math. Comp., 65:373–381, 1996.

[7] J. L. Selfridge C. Pomerance and Jr S. S. Wagstaff. *The Pseudoprimes to* $25 \cdot 10^9$. Mathematics of Computation, 35, pp 1003- 1026, 1980.

[8] Pomerance Damgård, Landrock. *Average Case Error Estimate for the strong probable prime test*. Mathematics of Computation, Vol 61, Number 203, 1993.

[9] J. Gilchrist. *Pseudoprime enumeration with probabilistic primality tests.* ttp://gilchrist. ca/jeff/factoring/pseudoprimes.html, 2013.

[10] N. Saxena M. Agrawal, N. Kayal. *Explicit bounds for primality testing and related problems*. Math. Comp. 55.191, pp. 355–380. DOI: 10.2307/2008811, 1990.

[11] N. Saxena M. Agrawal, N. Kayal. *PRIMES is in P*. Annals of Mathematics 160 (2): 781-793, 2004.

[12] G. L. Miller. *Riemann's hypothesis and tests for primality*. Jouurnal of Computer and System Sciences 13, no. 3, pp. 300–317., 1976.

[13] L. Monier. *Evaluation and comparison of two efficient probabilstic primality testing algorithms*. Theoretical Computer Science Volume 12, Issue 1, September 1980, Pages 97-108, 1978.

[14] L. Monier. *Evaluation and comparison of two efficient probabilistic primality testing algorithms*. Theoret. Comput. Sci., 12:97–108, 1980.

[15] C. Pomerance. *On the Distribution of Pseudoprimes*. Math. Comp. 37, 587-593, 1981.

[16] C. Pomerance. *A new lower bound for the pseudoprime counting function*. J. Math. 26, 4-9, 1982.

[17] C. Pomerance. *Are there counter-examples to the Baillie-PSW primality test.* Dopo Le Parole aangeboden aan Dr. A. K. Lenstra., 1984.

[18] S.S. Wagstaff Jr. R. Baillie. *Lucas pseudoprimes*. Math. Comp. 35, 1391-1417, 1980.

[19] M. Rabin. *Probabilistic algorithms*. : Traub JF (ed) Algorithms and complexity: new directions and recent results. Academic Press, New York, 1976.

[20] M. Rabin. *Probabilistic algorithm for testing primality*. J. Number Theory, 12:128–138, 1980.

[21] H. Riesel and R. C. Vaughan. *On the sums of primes*. Ark. Mat., 21(1):46–74, 1981.

[22] C. Pomerance S. H. Kim. *The probability that a random probable prime is composite*. Math. Comp. 53, 721–741., 1989.

[23] R. Solovay and V. Strassen. *fast Monte-Carlo test for primality*. SIAM J Comput 6(1):84-85, 1977.

[24] Mak Trifkovic. *Algebraic Theory of Quadratic Numbers*. Springer Science+Business Media New York, ISBN 978-1-4614-7716-7, 2013.

[25] Andrew Granville W. R. Alford and Carl Pomerance. *There are infinitely many Carmichael numbers*. In: Ann. of Math. (2) 139.3 (1994), pp. 703–722. DOI: 10.2307/ 2118576, 1994.

[26] H.C. Williams. *On numbers analogous to the Carmichael numbers*. Canad. Math. Bull. 20, 133-143, 1977.