TESTING FOR COMPOSITENESS

MIGUEL AMEZOLA

B.S. Mathematics
Pacific Lutheran University
Advisors: Dr. Tom Edgar

Abstract. The problem of distinguishing composite numbers from primes is an important task for mathematicians and many others. We explore two methods for performing such a binary classification of the integers. We begin with the mathematical formulation of a classic test for compositeness known as the Miller-Rabin test. This involves the exposition of concepts from number theory, such as divisibility and modular arithmetic, and algebraic structures like monoids, groups, and rings. We follow this with an evaluation of Miller-Rabin's performance in comparison with a modern machine learning algorithm called Support Vector Machine.

E-mail address: math@miguelamezola.com.
Date: May 19, 2017.

Contents

1. Introduction

Johann Carl Friedrich Gauss once wrote that "the problem of distinguishing prime numbers from composite numbers ... is known to be one of the most important and useful in arithmetic ... Further, the dignity of the science itself seems to require that every possible means be explored for the solution of a problem so elegant and so celebrated."[1] In this spirit, we will explore two methods for the binary classification of integers as either prime or composite.

We formally define these mutually exclusive classes of integers as follows.

Definition 1.0.1 (Prime)**.** Let $p \in \mathbb{Z}$, $p > 1$. Then p is prime if and only if for every $a, b \in \mathbb{Z}$, $p = ab$ implies $a = 1$ or $b = 1$. [2]

Definition 1.0.2 (Composite)**.** Let $n \in \mathbb{Z}, n > 1$. Then n is composite if and only if there exists $a, b \in \mathbb{Z}$ such that $n = ab, 1 < a, b < n$. [2]

Of course, all even integers greater than 2 are composite since they are all multiples of 2. Likewise, every multiple of 5 or 10 is composite. For relatively small integers, the existence of $a, b \in \mathbb{Z}$ such that $n = ab$ with $1 < a, b < n$ is easy to establish. For example, 63 is composite since $63 = 9 \cdot 7$ and 1023 is composite since $1023 = 341 \cdot 3$.

We notice that both 63 and 1023 have factors less than their squared roots; that is, $7 < \sqrt{63} \approx 7.94$ and $3 < \sqrt{1023} \approx 31.98$. Suppose that this is not true for some composite number $n = ab$. Then $a, b \geq \sqrt{n}$. But this leads to a contradiction, $n = ab > \sqrt{n} \cdot \sqrt{n} = n$ but n cannot be greater than itself. Thus, either $a < \sqrt{n}$ or $b < \sqrt{n}$.

This fact is used in a common test for compositeness called trial division. For instance, suppose that we would like to know whether or not 201 is composite. We know that $\sqrt{201} \approx 14.18$ and so we only have to check if 201 is a multiple of $2, 3, 4, \ldots, 14$. Since 201 is not even, we know that it is not a multiple of 2 and so we proceed with the next integer. We then discover, to our delight, that $201 = 3 \cdot 67$. Hence, 201 is composite.

Unfortunately, trial division is a laborious task for large integers like 154,320,140,719,313. The square root of this number is 12,422,566. That means that we would have to check whether or not any integer between 2 and 12,422,566 is a multiple of 154,320,140,719,313. Even if we noticed that every integer greater than or equal to 2 has a prime factor, and that any factor of a or b is also a factor of $ab = n$, and so every composite number n has a prime factor p such that $p < \sqrt{n}$, we would still have to perform trial division for 154,320,140,719,313 with every prime between 2 and its first prime factor 349!

For this reason, we would like to explore more efficient compositeness tests. First, we will consider a classic algorithm known as the Miller-Rabin test, followed by a method that uses a modern machine learning algorithm called a Support Vector Machine. However, before we are ready to do so, we must discuss the concepts in number theory and abstract algebra that form the foundation of the Miller-Rabin test.

2. Background for the Miller-Rabin Test

2.1. **Divisibility.**

Divisibility is a key concept in number theory. The Miller-Rabin test uses two key facts related to divisibility, namely,

(1) two integers with a greatest common divisor equal to 1 are relatively prime, and
(2) the greatest common divisor of two integers is the smallest possible linear combination of those integers.

Both facts rely on the following definition.

Definition 2.1.1 (Divide)**.** Let $a, d \in \mathbb{Z}$. We say that d divides a if there exists $q \in \mathbb{Z}$ such that $a = qd$. We express this in symbols as $d \mid a$ (which is read "d divides a"). We refer to d as a divisor of a. [2]

In our discussion regarding composite numbers, we concluded that 63 is composite since $63 = 9 \cdot 7$ and 1023 is composite since $1023 = 341 \cdot 3$. Thus, according to this definition, we can say that 7 divides 63 and 341 divides 1023. However, like most composite numbers, 63 and 1023 have more divisors, i.e. $1, 3, 7, 9, 21, 63$ divide 63 and $1, 3, 11, 31, 33, 93, 341, 1023$ divide 1023. Notice that the greatest number in both lists is 3. This is their greatest common divisor.

Definition 2.1.2 (Greatest common divisor)**.** Let $a, b \in \mathbb{Z}$ be nonzero. The greatest common divisor of a and b is the largest $d \in \mathbb{N}$ for which $d \mid a$ and $d \mid b$ and is denoted $d = \gcd(a, b)$.

However, the greatest common divisor of some integers is 1. For instance, the divisors of 25 are $1, 5, 15$ and the divisors of 28 are $1, 2, 4, 7, 14, 28$. The largest number that appears in both lists is 1, and so $\gcd(25, 28) = 1$. Such numbers are said to be relatively prime.

Definition 2.1.3 (Relatively prime)**.** Let $a, b \in \mathbb{Z}$. We say a and b are relatively prime if $\gcd(a, b) = 1$.

Proposition 2.1.1. *Let $a, p \in \mathbb{Z}$ such that p is prime and $p \nmid a$, then $\gcd(a, p) = 1$.*

Proof. We will use contradiction. Let $a, p \in \mathbb{Z}$ such that p is prime and $p \nmid a$. Suppose $\gcd(a, p) \neq 1$, then $a = kd$ and $p = ld$ for some $d, k, l \in \mathbb{Z}$ with $d \neq 1$. Since p is prime and $d \neq 1$, then $l = 1$ and $d = p$ by definition 1.0.1. But $a = kd = kp$ if and only if $p \mid a$, contradicting the fact that $p \nmid a$. Therefore, we must conclude that $\gcd(a, p) = 1$. □

Hence, two integers with a greatest common divisor equal to 1 are relatively prime. In particular, a prime number is relatively prime to any integer that it does not divide. Lastly, we will show that the greatest common divisor of two integers is the smallest possible linear combination of those integers.

Lemma 2.1.1 (Linear combination)**.** *Let $d, m, n, x, y \in \mathbb{Z}$. If $d \mid x$ and $d \mid y$, then $d \mid mx + ny$.*

Proof. Let $d, m, n, x, y \in \mathbb{Z}$ such that $d \mid x$ and $d \mid y$. Since $d \mid x$, then there exists $k \in \mathbb{Z}$ such that $x = kd$. Multiplying by m, we have $mx = mkd$. Since $d \mid y$, then there exists $l \in \mathbb{Z}$ such that $y = ld$. Multiplying by n, we have $ny = nld$. Adding both results yields

$$mx + ny = mkd + nld = (mk + nl)d.$$

Therefore, since $(mk + nl) \in \mathbb{Z}$, we know $d \mid mx + ny$. □

Proposition 2.1.2. *Let $a, b \in \mathbb{Z}$. Then greatest common divisor of a and b is equal to the smallest positive linear combination of a and b.*

Proof. We will use contradiction. Let $a, b \in \mathbb{Z}$, and suppose $e = ax + by$ is the smallest positive linear combination of a and b. Let $d = \gcd(a, b)$. Then $d \mid a$ and $d \mid b$ implies that $d \mid ax + ay$ by lemma 2.1.1. Since $d \mid ax + by$, then $d \mid e$, and so $d \leq e$.

Now suppose $e \nmid a$. Thus, $a = qe + r$ for some $q \in \mathbb{Z}$ with $0 < r < e$. Then

$$r = a - qe = a - q(ax - by) = a - qax - qby = a(1 - qx) - b(qy).$$

This means that r is a positive linear combination that is less than e, contradicting the fact that e is the smallest positive linear combination of a and b. Hence, $e \mid a$, and a similar argument can be used to show that $e \mid b$ as well. Thus, by lemma 2.1.1, $e \mid d$ and $e \mid ax + by$, and so $e \leq d$.

Since we have $d \leq e$ and $e \leq d$, it must be the case that $e = d$. Therefore, the greatest common divisor of a and b is equal to the smallest positive linear combination of a and b. □

2.2. **The Set of All Congruence Classes.** Divisibility can be used to partition the set of all integers into a finite number of classes. These classes are groups of integers with a common property related to the following definition.

Definition 2.2.1 (Congruent)**.** Let $a, b, n \in \mathbb{Z}$ with $n > 0$. We say that a is congruent to b modulo n if $n \mid (a - b)$, denoted $a \equiv b \pmod{n}$. [2]

Proposition 2.2.1. *Let $a, b, n \in \mathbb{Z}$. Then the following conditions are all equivalent.*

(1) $a = b + kn$ for some $k \in \mathbb{Z}$;
(2) $n \mid a - b$;
(3) $a \equiv b \pmod{n}$.

Proof. Let $a, b, n \in \mathbb{Z}$. First suppose condition (1). Then $a = b + kn$ implies $a - b = kn$. By the definition of divide, it follows that $n \mid a - b$. Thus, condition (1) implies condition (2).

Now suppose condition (2). Then, by the definition of congruent, $n \mid a - b$ implies $a \equiv b \pmod{n}$. Therefore, condition (2) implies condition (3).

Finally, we would like to show that condition (3) implies condition (1). So, we assume $a \equiv b \pmod{n}$. Then, by the definition of congruent, we know $n \mid a - b$. If follows that there exists $k \in \mathbb{Z}$ such that $a - b = kn$ by the definition of divide. Therefore, condition (3) implies condition (1). □

Example 2.2.1. Let $a = 34$, let $b = 144$, and let $n = 10$. Since $34 - 144 = -110 = -11 \cdot 10$, we have

(1) $34 = 144 + (-11) \cdot 10$,
(2) $10 \mid 34 - 144$, and
(3) $34 \equiv 144 \pmod{10}$

as expected.

However, 34 is not the only integer that is congruent to 4 modulo 10. In fact, every integer 4 more than a multiple of 10 also shares this property. This set of integers is an example of what we will call a congruence class.

Definition 2.2.2 (Congruence Class)**.** Let $a, n \in \mathbb{Z}$ with $n > 0$. We define the congruence class of a modulo n as the set of all integers congruent to a modulo n; that is,

$$\bar{a} := \{x \in \mathbb{Z} : x \equiv a \pmod{n}\}.$$

Each modulus partitions the entire set of integers into finitely many congruence classes.

Example 2.2.2. Let $n = 3$. Then $\bar{0} \in \mathbb{Z}_3$ contains all the multiples of 3, $\bar{1} \in \mathbb{Z}_3$ contains all the integers that are 1 more than a multiple of 3, and $\bar{2} \in \mathbb{Z}_3$ contains all the integers that are 2 more than an multiple of 3. However, integers that are congruent to 4 modulo 3 are equal to $1 + k \cdot 3$ according to proposition 2.2.1, which means that they are also 1 more than a multiple of 3. Thus, $\bar{4} = \bar{1}$ under this modulus. Hence, this modulus partitions the integers into exactly 3 congruence classes, namely, $\bar{0}, \bar{1}$, and $\bar{2}$.

Definition 2.2.3 ($\mathbb{Z}_n$)**.** Let $n > 0$ be any integer. We define $\mathbb{Z}_n$ to be the set of all congruence classes modulo n, i.e.

$$\mathbb{Z}_n := \{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{n-1}\}.$$

Proposition 2.2.2 ($\mathbb{Z}_n \neq \emptyset$)**.** *Let $n > 0$ be any integer. Then $\mathbb{Z}_n \neq \emptyset$.*

Proof. Let $n > 0$ be any integer. Then for $k = 1 \in \mathbb{Z}$, $n = 0 + kn$ implies $n \equiv 0 \pmod{n}$ by proposition 2.2.1. By the definition of congruence classes, it follows that $n \equiv 0 \pmod{n}$ is a member of $\bar{0} \in \mathbb{Z}_n$. Therefore, $\mathbb{Z}_n \neq \emptyset$. □

TABLE 1. Group-like algebraic structures consisting of associative binary operations.

	Identity	**Inverses**	**Commutativity**
Monoid	Required	Unneeded	Unneeded
Group	Required	Required	Unneeded
Abelian Group	Required	Required	Required

2.3. **The Algebraic Structure of $\mathbb{Z}_n$.** For the Miller-Rabin test, we need to define a multiplication and an addition on $\mathbb{Z}_n$. Furthermore, we must understand the structure of $\mathbb{Z}_n$ under these operations.

Definition 2.3.1 (Binary operation)**.** Let S be a set. We define a binary operation on S to be a function $f : S \times S \to S$ that assigns to each pair $(a, b) \in S \times S$ a unique element $a \circ b \in S$.

A binary operation $\circ$ with the property that $(a \circ b) \circ c = a \circ (b \circ c)$ for all $a, b, c \in S$ is called associative.

We now define two operations on $\mathbb{Z}_n$ that combine two elements to produce a third.

Definition 2.3.2 (Addition on $\mathbb{Z}_n$)**.** Let $\bar{a}, \bar{b} \in \mathbb{Z}_n$. Then addition on $\mathbb{Z}_n$ is the function $f : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$ defined by

$$f[(\bar{a}, \bar{b})] := \overline{a + b},$$

and denoted

$$\bar{a} + \bar{b} := \overline{a + b}.$$

We refer to this addition as addition on $\mathbb{Z}_n$ or addition modulo n.

Definition 2.3.3 (Multiplication on $\mathbb{Z}_n$)**.** Let $\bar{a}, \bar{b} \in \mathbb{Z}_n$. Then multiplication on $\mathbb{Z}_n$ is the function $f : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$ defined by

$$f[(\bar{a}, \bar{b})] := \overline{a \cdot b},$$

and denoted

$$\bar{a} \cdot \bar{b} := \overline{a \cdot b}.$$

We refer to this multiplication as multiplication on $\mathbb{Z}_n$ or multiplication modulo n, and use the abbreviation ab for $a \cdot b$.

Lemma 2.3.1. *Addition and multiplication are closed and well-defined on $\mathbb{Z}_n$.*

Proof. First, we must show that addition and multiplication on $\mathbb{Z}_n$ are closed. By definition 2.3.2, we know that for any $\bar{a}, \bar{b} \in \mathbb{Z}_n$, we have $\bar{a} + \bar{b} = \overline{a + b} \in \mathbb{Z}_n$ and $\bar{a} \cdot \bar{b} = \overline{a \cdot b} \in \mathbb{Z}_n$. Thus, addition and multiplication are closed.

Now we will show that this addition is well-defined. Suppose $\bar{a}, \bar{b}, \bar{c}, \bar{d} \in \mathbb{Z}_n$, and that $\bar{a} = \bar{c}$ and $\bar{b} = \bar{d}$. Then, by proposition 2.2.1, $a - c = pn$ and $b - d = qn$ for some $p, q \in \mathbb{Z}$. Adding both equations, we have $a - c + b - d = pn + qn$, which can we rewritten as $(a + b) - (c + d) = (p + q)n$. This implies that $a + b \equiv c + d \pmod{n}$ and so $\overline{a + b} = \overline{c + d}$. Hence, this addition is well-defined.

On the other hand, if we multiply $a - c = pn$ by b and $b - d = qn$ by c, we have $ab - bc = bpn$ and $bc - cd = cqn$. Adding both equations tells us $ab - cd = ab - bc + bc - cd = bpn + cqn = (bp + cq)n$. Since $bp + cq \in \mathbb{Z}$, we have shown that $ab \equiv cd \pmod{n}$ and so $\overline{a \cdot b} = \overline{c \cdot d}$ as required. Thus, multiplication modulo n is well-defined. □

Table 1 lists three group-like algebraic structures: monoid, group, and abelian group. They each consist of a set together with a binary operation that satisfy a list of axioms. We will show that $\mathbb{Z}_n$ is a monoid under both operations defined above, and that $\mathbb{Z}_n$ forms an abelian group under addition. This will lead to the conclusion that, since this multiplication distributes over this addition, $\mathbb{Z}_n$ must also be a ring. We begin by formally defining a monoid.

Monoid.

Definition 2.3.4. A monoid is a set M together with a binary operation $\circ$ that satisfies the following axioms.

(1) The binary operation is associative. That is,

$$a \circ (b \circ c) = (a \circ b) \circ c$$

for all $a, b, c \in M$.

(2) There exists an element $e \in M$, called the identity element, such that for any element $a \in M$

$$e \circ a = a \circ e = a.$$

Proposition 2.3.1. *Let $n > 0$ be any integer. Then the set $\mathbb{Z}_n$ forms a monoid under addition modulo n.*

Proof. Let $n > 0$ be any integer. First, we recall that $\mathbb{Z}_n \neq \emptyset$ by proposition 2.2.2, and that addition is well-defined and closed binary operation according to lemma 2.3.1. Now, we must show that this operation is associative. So, let $\bar{a}, \bar{b}, \bar{c} \in \mathbb{Z}_n$. Then, by the associativity of addition on the integers, we have

$$(\bar{a} + \bar{b}) + \bar{c} = \overline{a + b} + \bar{c} = \overline{(a + b) + c} = \overline{a + (b + c)} = \bar{a} + \overline{b + c} = \bar{a} + (\bar{b} + \bar{c})$$

as required.

Next, we must show that $\mathbb{Z}_n$ contains an identity element under this operation. We saw in our proof of proposition 2.2.2 that $\bar{0} \in \mathbb{Z}_n$. Then $\bar{a} + \bar{0} = \overline{a + 0} = \bar{a} = \overline{0 + a} = \bar{0} + \bar{a}$; that is, $\bar{a} + \bar{0} = \bar{a} = \bar{0} + \bar{a}$. Hence, $\bar{0} \in \mathbb{Z}_n$ is an identity element under this addition for all $\bar{a} \in \mathbb{Z}_n$. Therefore, we conclude that the set $\mathbb{Z}_n$ forms a monoid under addition modulo n. □

Proposition 2.3.2. *Let $n > 0$ be any integer. Then the set $\mathbb{Z}_n$ forms a monoid under multiplication modulo n.*

Proof. Let $n > 0$ be any integer. Then $\mathbb{Z}_n$ is nonempty by proposition 2.2.2. Moreover, by lemma 2.3.1, we know that this multiplication is a closed and well-defined binary operation on this set. Now, we must show that this operation is associative. Let $\bar{a}, \bar{b}, \bar{c} \in \mathbb{Z}_n$. Then, by the associativity of multiplication on the integers, we have

$$(\bar{a} \cdot \bar{b}) \cdot \bar{c} = \overline{a \cdot b} \cdot \bar{c} = \overline{(a \cdot b) \cdot c} = \overline{a \cdot (b \cdot c)} = \bar{a} \cdot \overline{b \cdot c} = \bar{a} \cdot (\bar{b} \cdot \bar{c})$$

as required.

Lastly, we must prove that $\mathbb{Z}_n$ has an identity element under this multiplication. Since $(n+1) = 1 + kn$ for $k = 1 \in \mathbb{Z}$, we know $(n + 1) \equiv 1 \pmod{n}$ by proposition 2.2.1; which is an element in $\bar{1} \in \mathbb{Z}_n$. Thus, $\bar{1} \in \mathbb{Z}_n$. It follows, by the commutativity of integer multiplication, that

$$\bar{1} \cdot \bar{a} = \overline{1 \cdot a} = \bar{a} = \overline{a \cdot 1} = \bar{a} \cdot \bar{1};$$

that is, $\bar{1} \cdot \bar{a} = \bar{a} = \bar{a} \cdot \bar{1}$. Hence, $\bar{1} \in \mathbb{Z}_n$ is an identity element under this multiplication. Therefore, $\mathbb{Z}_n$ forms a monoid under multiplication modulo n. □

We will make use of exponential notation in computations related to the Miller-Rabin test.

Definition 2.3.5 (Exponential notation)**.** Let M be a monoid, let $a \in M$, and let $e \in M$ be the identity element. We first define $a^0 = e$. For $n \in \mathbb{N}$ with $n > 0$, we define

$$a^n = \underbrace{a \circ a \circ \cdots \circ a}_{n \text{ times}}$$

and

$$a^{-n} = \underbrace{a^{-1} \circ a^{-1} \circ \cdots \circ a^{-1}}_{n \text{ times}}.$$

Proposition 2.3.3. *Let M be a monoid, and let $a \in M$. Then, for all $m, n \in \mathbb{Z}$,*

(1) $a^m \circ a^n = a^{m+n}$, and
(2) $(a^m)^n = a^{mn}$.

Proof. Let M be a monoid, let $a \in M$, and let $m, n \in \mathbb{Z}$. Then, by the definition of exponential notation and the associativity of the group operation, we have

$$\begin{aligned} a^m \circ a^n &= \underbrace{(a \circ a \circ \cdots \circ a)}_{m \text{ times}} \circ \underbrace{(a \circ a \circ \cdots \circ a)}_{n \text{ times}} \\ &= \underbrace{a \circ a \circ \cdots \circ a \circ a \circ a \circ \cdots \circ a}_{m+n \text{ times}} \\ &= a^{m+n} \end{aligned}$$

and

$$\begin{aligned} (a^m)^n &= \underbrace{a^m \circ a^m \circ \cdots \circ a^m}_{n \text{ times}} \\ &= \underbrace{\overbrace{(a \circ a \circ \cdots \circ a)}^{m \text{ times}} \circ \overbrace{(a \circ a \circ \cdots \circ a)}^{m \text{ times}} \circ \cdots \circ \overbrace{(a \circ a \circ \cdots \circ a)}^{m \text{ times}}}_{n \text{ times}} \\ &= \underbrace{a \circ a \circ \cdots \circ a \circ a \circ a \circ \cdots \circ a \circ \cdots \circ a \circ a \circ \cdots \circ a}_{mn \text{ times}} \\ &= a^{mn}. \end{aligned}$$

□

Now that we are convinced that $\mathbb{Z}_n$ forms a monoid under both operations, we will continue to explore the structure of this set under modular addition.

Abelian Group.

Definition 2.3.6 (Group)**.** We define a group G to be a monoid that contains an inverse element for each element $a \in G$, denoted by a^{-1}, such that

$$a \circ a^{-1} = a^{-1} \circ a = e.$$

A group with the property that $a \circ b = b \circ a$ for $a, b \in G$ is called abelian.

Proposition 2.3.4. *Let $n > 0$ be any integer. Then the set $\mathbb{Z}_n$ forms an abelian group under addition modulo n.*

Proof. Let $n > 0$ be any integer. According to proposition 2.3.1, $\mathbb{Z}_n$ is a monoid under addition modulo n. Now we must show that the set $\mathbb{Z}_n$ contains an inverse element for each of its members. Let $\bar{a} \in \mathbb{Z}_n$. Then, since $(n - a) = (n - a) + kn$ for $k = 0 \in \mathbb{Z}$, it follows that $(n - a) \equiv (n - a) \pmod{n}$. Hence, $\overline{n - a} \in Z_n$ by the definition of congruence classes. Adding both of these elements,

we have

$$\begin{aligned}\bar{a} + \overline{n-a} &= \overline{a+n-a} \\ &= \overline{a+(-a)+n} \\ &= \bar{n} \\ &= \overline{n+a-a} \\ &= \overline{n+(-a)+a} \\ &= \overline{n-a} + \bar{a}\end{aligned}$$

by the commutativity of addition on the integers. Thus, $\bar{a} + \overline{n-a} = \overline{n-a} + \bar{a} = \bar{a}$, and so $\mathbb{Z}_n$ is a group under this operation.

To show that it is abelian, let $\bar{b} \in \mathbb{Z}_n$. Again by the commutativity of integer addition, we have

$$\bar{a} + \bar{b} = \overline{a+b} = \overline{b+a} = \bar{b} + \bar{a}.$$

Since both of these elements were arbitrary, we conclude that this must be true for all elements in $\mathbb{Z}_n$. Therefore, $\mathbb{Z}_n$ forms an abelian group under addition modulo n. □

Indeed, we have learned much about the structure of $\mathbb{Z}_n$. We now know that $\mathbb{Z}_n$ forms a monoid under multiplication and an abelian group under addition. For the sake of conciseness, we will combine these concepts in order to form a single algebraic structure called a ring.

Ring.

Definition 2.3.7 (Ring)**.** Let R be an abelian group. Then R is a ring if it forms a monoid under a second binary operation $\circ$ that distributes over the group operation.

Proposition 2.3.5. *Let $n > 0$ be any integer. Then $\mathbb{Z}_n$ forms a ring under addition modulo n and multiplication modulo n.*

Proof. Let $n > 0$ be any integer. Then, by proposition 2.3.4, $\mathbb{Z}_n$ forms an abelian group under addition modulo n. Moreover, we know $\mathbb{Z}_n$ forms a monoid under multiplication modulo n by proposition 2.3.2.

So, all we have left to show is that this multiplication distributes over this addition. Let $\bar{a}, \bar{b}, \bar{c} \in \mathbb{Z}_n$. Since multiplication distributes over addition on the integers, we have

$$\begin{aligned}\bar{a} \cdot (\bar{b} + \bar{c}) &= \bar{a} \cdot \overline{b+c} \\ &= \overline{a \cdot (b+c)} \\ &= \overline{a \cdot b + ac} \\ &= \overline{a \cdot b} + \overline{a \cdot c} \\ &= \bar{a} \cdot \bar{b} + \bar{a} \cdot \bar{c}\end{aligned}$$

for left distribution, and

$$\begin{aligned}(\bar{a} + \bar{b}) \cdot \bar{c} &= \overline{a+b} \cdot \bar{c} \\ &= \overline{(a+b) \cdot c} \\ &= \overline{a \cdot c + b \cdot c} \\ &= \overline{a \cdot c} + \overline{b \cdot c} \\ &= \bar{a} \cdot \bar{c} + \bar{b} \cdot \bar{c}\end{aligned}$$

for right distribution. Thus, this multiplication distributes over addition modulo n. Therefore, the set $\mathbb{Z}_n$ together with this addition and this multiplication forms a ring. □

2.4. **Prime Moduli.** Table 2 and table 3 show multiplication in $\mathbb{Z}_7$ and $\mathbb{Z}_8$, respectively. We notice that in $\mathbb{Z}_7$, each row and each column contain the identity element exactly once. This indicates that each element in $\mathbb{Z}_7$ has a unique multiplicative inverse. On the other hand, $\bar{1}$ only appears in half of the rows in the table for $\mathbb{Z}_8$. Thus, elements like $\bar{2}$ do not have multiplicative inverses in $\mathbb{Z}_8$.

TABLE 2. Multiplication in $\mathbb{Z}_7$

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{6}$	$\bar{1}$	$\bar{3}$	$\bar{5}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{6}$	$\bar{2}$	$\bar{5}$	$\bar{1}$	$\bar{4}$
$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{1}$	$\bar{5}$	$\bar{2}$	$\bar{6}$	$\bar{3}$
$\bar{5}$	$\bar{0}$	$\bar{5}$	$\bar{3}$	$\bar{1}$	$\bar{6}$	$\bar{4}$	$\bar{2}$
$\bar{6}$	$\bar{0}$	$\bar{6}$	$\bar{5}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

TABLE 3. Multiplication in $\mathbb{Z}_8$

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{6}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{6}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{6}$	$\bar{1}$	$\bar{4}$	$\bar{7}$	$\bar{2}$	$\bar{5}$
$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{0}$	$\bar{4}$
$\bar{5}$	$\bar{0}$	$\bar{5}$	$\bar{2}$	$\bar{7}$	$\bar{4}$	$\bar{1}$	$\bar{6}$	$\bar{3}$
$\bar{6}$	$\bar{0}$	$\bar{6}$	$\bar{4}$	$\bar{2}$	$\bar{0}$	$\bar{6}$	$\bar{4}$	$\bar{2}$
$\bar{7}$	$\bar{0}$	$\bar{7}$	$\bar{6}$	$\bar{5}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Since 7 is prime and 8 is composite, we suspect that this difference in their multiplication tables has something to do with the primality of the modulus.

Lemma 2.4.1. *Let p be prime. Then for each $\bar{a} \in \mathbb{Z}_p$ with $\bar{a} \neq \bar{0}$, there exists a unique multiplicative inverse.*

Proof. Let p be prime, and let $\bar{a} \in \mathbb{Z}_p$ with $\bar{a} \neq \bar{0}$. Then, by proposition 2.1.1, we have $\gcd(a, p) = 1$. Thus, $ax + py = 1$ according to proposition 2.1.2. It follows that $ax = 1 + (-y)p$; that is, $ax \equiv 1 \pmod{p}$ by proposition 2.2.1. Hence, $\overline{a \cdot x} = \bar{a} \cdot \bar{x} = \bar{1}$, and so every nonzero element in $\mathbb{Z}_p$ has a multiplicative inverse

Now we must show uniqueness. Suppose that both $\bar{x}, \bar{y} \in \mathbb{Z}_p$ are multiplicative inverses of $\bar{a}$. Then $\bar{a} \cdot \bar{x} = \bar{1}$ and $\bar{a} \cdot \bar{y} = \bar{1}$. Thus,

$$\begin{aligned} \bar{a} \cdot \bar{x} &= \bar{a} \cdot \bar{y} \\ \bar{a}^{-1} \cdot \bar{a} \cdot \bar{x} &= \bar{a}^{-1} \cdot \bar{a} \cdot \bar{y} \\ \bar{x} &= \bar{y} \end{aligned}$$

Therefore, for each $\bar{a} \in \mathbb{Z}_p$ with $\bar{a} \neq \bar{0}$, there exists a unique multiplicative inverse. □

This result is important. We will use it to prove two claims related to prime moduli:

(1) $\mathbb{Z}_p$ has the zero product property, and
(2) $\bar{a}^{p-1} = \bar{1}$ for all nonzero $\bar{a} \in \mathbb{Z}_p$.

For both of these proofs, we need to know whether or not multiplication on $\mathbb{Z}_n$ is commutative.

Lemma 2.4.2. *Let $n > 1$, and let $\bar{a}, \bar{b} \in \mathbb{Z}_n$. Then $\bar{a} \cdot \bar{b} = \bar{b} \cdot \bar{a}$.*

Proof. Let $n > 0$, and let $\bar{a}, \bar{b} \in \mathbb{Z}_n$. Then, by definition 2.3.3, we have $\bar{a} \cdot \bar{b} = \overline{a \cdot b}$. Since integer multiplication is commutative, $\overline{a \cdot b} = \overline{b \cdot a}$. Finally, by definition 2.3.3 again, we have $\overline{b \cdot a} = \bar{b} \cdot \bar{a}$. Therefore, $\bar{a} \cdot \bar{b} = \bar{b} \cdot \bar{a}$. □

Zero Product Property.

Theorem 2.4.1 (Zero product property)**.** *Let p be prime. If $\bar{a} \cdot \bar{b} = \bar{0}$, then either $\bar{a} = \bar{0}$ or $\bar{b} = \bar{0}$ for all $\bar{a}, \bar{b} \in \mathbb{Z}_p$.*

Proof. Let p be prime, and let $\bar{a}, \bar{b} \in \mathbb{Z}_p$ such that $\bar{a} \cdot \bar{b} = \bar{0}$. If both $\bar{a} = \bar{0}$ and $\bar{b} = \bar{0}$, then

$$\bar{a} \cdot \bar{b} = \bar{0} \cdot \bar{0} = \overline{0 \cdot 0} = \bar{0}.$$

Now let's suppose that $\bar{a}, \bar{b}$ are both not zero. It follows from lemma 2.4.1 that every nonzero element in $\mathbb{Z}_p$ has a multiplicative inverse. Since multiplication on $\mathbb{Z}_p$ is commutative according to lemma 2.4.2, we can assume without loss of generality that $\bar{a} \neq \bar{0}$. Then

$$\begin{aligned} \bar{a} \cdot \bar{b} &= \bar{0} \\ \bar{a}^{-1} \cdot \bar{a} \cdot \bar{b} &= \bar{a}^{-1} \cdot \bar{0} \\ \bar{1} \cdot \bar{b} &= \bar{0} \\ \bar{b} &= \bar{0}. \end{aligned}$$

Therefore, if $\bar{a} \cdot \bar{b} = \bar{0}$, then either $\bar{a} = \bar{0}$ or $\bar{b} = \bar{0}$ for all $\bar{a}, \bar{b} \in \mathbb{Z}_p$.

□

Fermat's Little Theorem. Fermat's Little Theorem states that, if p is prime, then $\bar{a}^{p-1} = \bar{1}$ for all nonzero $\bar{a} \in \mathbb{Z}_p$. Before we can prove this, we must show that multiplication on $\mathbb{Z}_n$ is commutative and that, when p is prime, $\bar{1} \in \mathbb{Z}_p$ and $\overline{p-1} \in \mathbb{Z}_p$ are their own inverses.

Lemma 2.4.3. *Let p be prime. Then $\bar{a} \in \mathbb{Z}_p$ is its own multiplicative inverse if and only if $\bar{a} = \bar{1}$ or $\bar{a} = \overline{p-1}$.*

Proof. Let $p \in \mathbb{N}$ be prime, and let $\bar{a} \in \mathbb{Z}_p$ be its own multiplicative inverse. Then, by definition 2.3.5, we have $\bar{a} \cdot \bar{a} = \bar{a}^2 = \bar{1}$. Furthermore, since $\mathbb{Z}_p$ is a ring, multiplication distributes over addition, and so $\bar{a}^2 - \bar{1} = (\bar{a} + \bar{1})(\bar{a} - \bar{1}) = \bar{0}$. If follows, by theorem 2.4.1, that since $(\bar{a} + \bar{1})(\bar{a} - \bar{1}) = \bar{0}$, then either $(\bar{a} + \bar{1}) = \bar{0}$ or $(\bar{a} - \bar{1}) = \bar{0}$. We will now consider both cases.

Case 1. If $(\bar{a} + \bar{1}) = \bar{0}$, then $\bar{a} = -\bar{1} = \overline{p-1}$.

Case 2. If $(\bar{a} - \bar{1}) = \bar{0}$, then $\bar{a} = \bar{1}$.

Thus, $\bar{a} = \bar{1}$ or $\bar{a} = \overline{p-1}$.

Conversely, for $\bar{1} \in \mathbb{Z}_p$, we have $\bar{1} \cdot \bar{1} = \overline{1 \cdot 1} = \bar{1}$ as required. For $\overline{p-1} \in \mathbb{Z}_p$, we have

$$\overline{p-1} \cdot \overline{p-1} = \overline{(p-1)(p-1)} = \overline{p^2 + (-2)p + 1} = \overline{p^2} + \overline{-2} \cdot \bar{p} + \bar{1}.$$

That is, $\overline{p-1} \cdot \overline{p-1} = \overline{p^2} + \overline{-2} \cdot \bar{p} + \bar{1}$. Since $p = 0 + kp$ for some $k \in \mathbb{Z}$, we know $\bar{p} = \bar{0}$ according to proposition 2.2.1. Hence,

$$\overline{p-1} \cdot \overline{p-1} = \overline{p^2} + \overline{-2} \cdot \bar{p} + \bar{1} = \bar{0} + \overline{-2} \cdot \bar{0} + \bar{1} = \bar{0} + \overline{-2 \cdot 0} + \bar{1} = \bar{1}.$$

Thus, $\overline{p-1} \cdot \overline{p-1} = \bar{1}$. □

Theorem 2.4.2 (Fermat's Little Theorem)**.** *Let p be prime, and let $\bar{a} \in \mathbb{Z}_p, \bar{a} \neq \bar{0}$. Then $\bar{a}^{p-1} = \bar{1}$.* [2]

Proof. Let p be prime, and let $\bar{a} \in \mathbb{Z}_p, \bar{a} \neq \bar{0}$. By lemma 2.4.1, we know that $\mathbb{Z}_p$ contains a unique inverse for each of its elements. Furthermore, $\bar{1}^{-1} = \bar{1}$ and $\overline{p-1}^{-1} = \overline{p-1}$ by lemma 2.4.3, and these are the only self-inverse elements. Thus, $\bar{1} \cdot \bar{2} \cdot \bar{3} \cdots \overline{p-1} = \bar{1} \cdot \overline{p-1} = \overline{p-1}$. Then

$$(\bar{a} \cdot \bar{1})(\bar{a} \cdot \bar{2}) \cdots (\bar{a} \cdot \overline{p-1}) = \underbrace{\bar{a} \cdot \bar{a} \cdots \bar{a}}_{p-1 \text{ times}} \cdot \bar{1} \cdot \bar{2} \cdots \bar{a} \cdots \bar{a}^{-1} \cdots \overline{p-1} = \bar{a}^{p-1} \cdot \overline{p-1}.$$

Moreover, since multiplication by nonzero $\bar{a} \in \mathbb{Z}_p$ is a one-to-one correspondence by lemma 2.4.1, we know that each product is equal to a unique element in $\mathbb{Z}_p$. Thus, $(\bar{a} \cdot \bar{1})(\bar{a} \cdot \bar{2}) \cdots (\bar{a} \cdot \overline{p-1}) = \bar{1} \cdot \bar{2} \cdots \overline{p-1}$, where the right-hand side is some permutation of the elements in $\mathbb{Z}_p$. Hence,

$$\bar{a}^{p-1} = \bar{a}^{p-1} \cdot \bar{1} = \bar{a}^{p-1} \cdot \overline{p-1} \cdot \overline{p-1} = (\bar{1} \cdot \bar{2} \cdot \ldots \cdot \overline{p-1})(\cdot \overline{p-1}) = (\overline{p-1})(\overline{p-1}) = 1.$$

Therefore, if p is prime, then $\bar{a}^{p-1} = \bar{1}$ for all $\bar{a} \in \mathbb{Z}_p, \bar{a} \neq \bar{0}$. □

Fermat's Little Theorem is the cornerstone of the Miller-Rabin test.

3. A Probabilistic Test for Compositeness

3.1. **Miller-Rabin Test.** The Miller-Rabin test uses three important results, namely,

(1) Fermat's Little Theorem,
(2) $\mathbb{Z}_n$ is a ring, and
(3) $\mathbb{Z}_n$ has the zero product property when n is prime.

Let $p > 2$ be prime. Fermat's Little Theorem tells us that when we choose a nonzero $\bar{a} \in \mathbb{Z}_p$, then $\bar{a}^{p-1} = \bar{1}$ or equivalently $\bar{a}^{p-1} - \bar{1} = \bar{0}$. Since $\mathbb{Z}_p$ is a ring, we know that multiplication distributes over addition. Thus, we can factor $\bar{0} = \bar{a}^{p-1} - \bar{1} = (\bar{a}^{\frac{p-1}{2}} - \bar{1})(\bar{a}^{\frac{p-1}{2}} + \bar{1})$. Recall that 2 is the only even prime. So, if p is a prime greater than 2, then $p-1$ must be even and so $2 \mid p-1$. Now we use the zero product property to conclude that if $(\bar{a}^{\frac{p-1}{2}} - \bar{1})(\bar{a}^{\frac{p-1}{2}} + \bar{1}) = \bar{0}$, then $(\bar{a}^{\frac{p-1}{2}} - \bar{1}) = \bar{0}$ or $(\bar{a}^{\frac{p-1}{2}} + \bar{1}) = \bar{0}$.

Example 3.1.1. Since 29 is prime, we know that $\bar{a}^{28} = \bar{1}$ for all $\bar{a} \in \mathbb{Z}_{29}$ by theorem 2.4.2. In other words, $\bar{a}^{28} - \bar{1} = \bar{0}$. Since $\mathbb{Z}_{29}$ is a field, this polynomial can be factored using the difference of squares. Then

$$\begin{aligned} \bar{a}^{28} - \bar{1} &= (\bar{a}^{14} + \bar{1})(\bar{a}^{14} - \bar{1}) \\ &= (\bar{a}^{14} + \bar{1})(\bar{a}^{7} + \bar{1})(\bar{a}^{7} - \bar{1}) \\ &= \bar{0}. \end{aligned}$$

Since 7 is odd, it is not possible to use the difference of squares to factor $(\bar{a}^7 - \bar{1})$ any further. By theorem 2.4.1, we know $(\bar{a}^{14} + \bar{1})(\bar{a}^7 + \bar{1})(\bar{a}^7 - \bar{1}) = \bar{0}$ implies that either $(\bar{a}^{14} + \bar{1}) = \bar{0}$ or $(\bar{a}^7 + \bar{1}) = \bar{0}$ or $(\bar{a}^7 - \bar{1}) = \bar{0}$. Even if we randomly select an $\bar{a} \in \mathbb{Z}_{29}$, we expect this to still be true. So, let $\bar{a} = \bar{7}$. Then we have

$$\begin{aligned} (\bar{7}^{14} + \bar{1}) &= \bar{2} \\ (\bar{7}^{7} + \bar{1}) &= \bar{2} \\ (\bar{7}^{7} - \bar{1}) &= \bar{0}, \end{aligned}$$

as we expected.

The Miller-Rabin test assumes that every odd $n > 1$ is prime and then performs the factorization described here. Then if at least one of the factors does not equal $\bar{0}$, we reach a contradiction and so have to conclude that n is composite. The following is a concise description of this method.

Algorithm 3.1.1 (Miller-Rabin Test for Compositeness)**.** *Let $n > 1$ be any odd integer. Then there exists an integer $k > 0$ such that 2^k is the largest power of two that divides $n-1$. If there exists $\bar{a} \in \mathbb{Z}_n$ such that*

$$\bar{a}^{\frac{n-1}{2^k}} \neq \bar{1}$$

and

$$\bar{a}^{\frac{n-1}{2^h}} \neq -\bar{1},$$

for all $h \in \mathbb{Z} : 1 \leq h \leq k$, then n is composite. In this case, the integer a is called a Miller-Rabin witness to the compositeness of n.

Example 3.1.2. We would like to use algorithm 3.1.1 to test the compositeness of 169. Since 2^3 is the largest power of two that divides 168, we must find an $\bar{a} \in \mathbb{Z}_{169}$ such that $\bar{a}^{\frac{168}{2^3}} \neq \bar{1}$ and $\bar{a}^{\frac{168}{2^h}} \neq -\bar{1}$ for all h, $h = 1, 2, 3$. So, we randomly choose $\overline{19} \in \mathbb{Z}_{169}$ and find that

$$\begin{aligned}\overline{19}^{\frac{168}{2^3}} &= \overline{70}\\ \overline{19}^{\frac{168}{2^2}} &= -\bar{1}\\ \overline{19}^{\frac{168}{2^1}} &= \bar{1}.\end{aligned}$$

Because $\overline{19}^{\frac{168}{2^2}} = -\bar{1}$, we cannot conclude that 169 is composite. So we randomly select a different $\bar{a} \in \mathbb{Z}_{169}$, namely $\bar{a} = \overline{145}$, and this time discover that

$$\begin{aligned}\overline{145}^{\frac{168}{2^3}} &= \overline{18}\\ \overline{145}^{\frac{168}{2^2}} &= \overline{155}\\ \overline{145}^{\frac{168}{2^1}} &= \overline{27}.\end{aligned}$$

Hence, 145 is a Miller-Rabin witness to the compositeness of 169 and we conclude that 169 is not prime.

3.2. **Effectiveness of the Miller-Rabin Test.** As seen in example 3.1.2, not every element in $\mathbb{Z}_n$ is a Miller-Rabin witness to the compositeness of n. In fact, in addition to $\overline{19}$, there are 11 more nonwitnesses in $\mathbb{Z}_{169}$, namely, $\bar{1}$, $\overline{22}$, $\overline{23}$, $\overline{70}$, $\overline{80}$, $\overline{89}$, $\overline{99}$, $\overline{146}$, $\overline{147}$, $\overline{150}$, and $\overline{168}$. Thus, if we randomly choose an element in $\mathbb{Z}_{169}$, the probability of choosing a Miller-Rabin witness is greater than 0.9 (see appendix A).

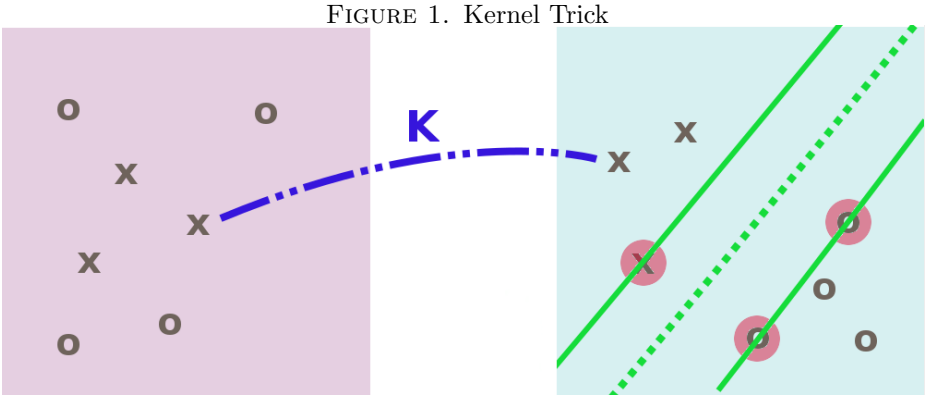
In general, the probability of randomly choosing a Miller-Rabin witnesses is 0.75, since, for any $n > 1$, at least 3 out of 4 elements in $\mathbb{Z}_n$ are witnesses [2]. Thus, by choosing random elements, we only have to repeat the Miller-Rabin test a relatively small number of times in order to be very confident of the compositeness of a number.

Example 3.2.1. At the outset, we concluded that trial division is a laborious task for large integers like 154,320,140,719,313. How does the Miller-Rabin test compare? We randomly choose $\overline{13} \in \mathbb{Z}_{154,320,140,719,313}$, and compute the exponents for the left-hand side (see appendix B). Then

$$\begin{aligned}\overline{13}^{77,160,070,359,656} &= \overline{124,576,103,502,502} \neq -\bar{1}\\ \overline{13}^{38,580,035,179,828} &= \overline{149,582,624,540,097} \neq -\bar{1}\\ \overline{13}^{19,290,017,589,914} &= \overline{130,308,615,937,868} \neq -\bar{1}\\ \overline{13}^{9,645,008,794,957} &= \overleftarrow{126,524,304,989,572} \neq -\bar{1}\\ \overline{13}^{9,645,008,794,957} &= \overleftarrow{126,524,304,989,572} \neq \bar{1}.\end{aligned}$$

Hence, $\overline{13}$ is a witness to the compositeness of $154,320,140,719,313$. On other hand, trial division would have to check 348 integers or 70 primes before coming to the same conclusion.

FIGURE 1. Kernel Trick



If we run the test 1 time, the probability of not finding a witness is $\frac{1}{4}$. In other words, let N be the number of times we repeat the test without finding a witness, then there is a $1 - \left(\frac{1}{4}\right)^N$ probability that the modulus is not composite and, therefore, prime. For example, consider 29,497,513,910,652,490,397. If we randomly choose $\overline{5,139,106,524} \in \mathbb{Z}_{29,497,513,910,652,490,397}$ and perform the Miller-Rabin test, we find that this element is not a witness. Hence, there is a probability of $1 - \frac{1}{4} = \frac{3}{4}$ that 29,497,513,910,652,490,397 is prime. Next, we randomly choose a different element, $\overline{5,432,093,847} \in \mathbb{Z}_{29,497,513,910,652,490,397}$, and discover that this element is not a witness either. Now there is a $1 - \left(\frac{1}{4}\right)^2 = 0.9375$ probability that this modulus is prime. We only have to repeat the test 5 more times before this probability exceeds 0.9999!

4. Support Vector Machine for Binary Classification

In reference to the Support Vector Machine (SVM) algorithm, Massachusetts Institute of Technology professor Patrick Winston once said, "this need to be in the tool bag of every civilized person." [6] Proposed by Vladamir N. Vapnik in the early 1990's, SVM is a "non-probabilistic binary linear classifier"[5] that is commonly used for text categorization, image classification, and hand-written character recognition. We would like to know how well SVM classifies integers as prime or composite.

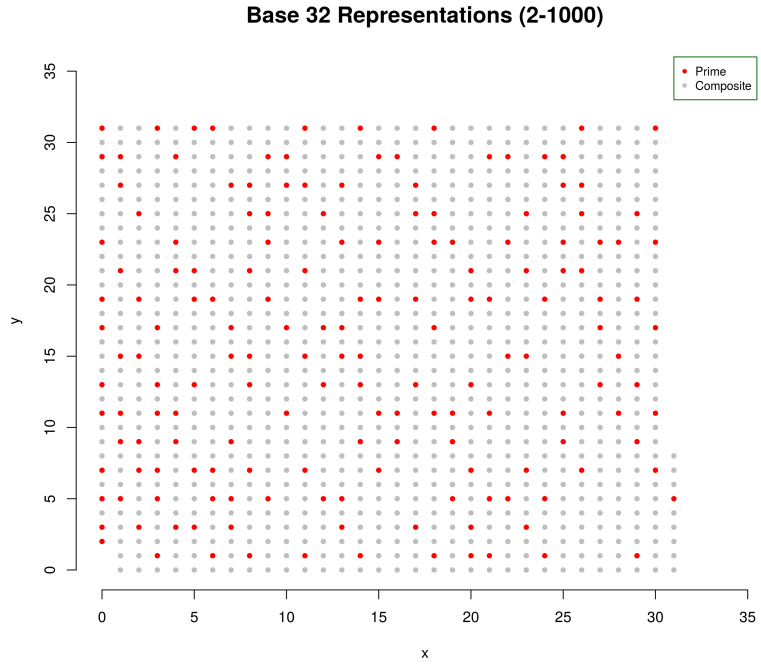
4.1. **Design.**

Linear Classification. SVM uses the so-called "Kernel trick" to separate points from different classes. Figure 1 shows two spaces with vectors of different classes. Since we cannot use a straight line (or hyperplane in higher dimensions) to separate the two classes in the left space, we use a nonlinear transformation K (commonly referred to as a Kernel function) to map these vectors into a higher dimensional space where the two classes are separable. Of course, this trick requires that we first assume that such a function exists.

Nonlinear Transformations. The following functions are commonly used for the nonlinear transformation K. [3]

- linear: $\langle x, x' \rangle$.
- polynomial: $(\gamma \langle x, x' + r \rangle^d$, with $\gamma, r \in \mathbb{R}$ and $d \in \mathbb{N}$.
- rbf: $\langle x, x' \rangle$, where $\gamma \in \mathbb{R}$ with $\gamma > 0$.
- sigmoid: $\langle x, x' \rangle$, with $r \in \mathbb{R}$.

FIGURE 2. Base 32 representations of integers between 2 and 1000 (inclusive).



Widest Street Approach. Once the classes are linearly separable, we take the "widest street approach" [6] to construct the optimal hyperplane. This means choosing the hyperplane that creates the largest margin possible between vectors of different classes. In general, "the larger the margin the lower the generalization error." [3] SVM uses a Lagrangian method for constrained optimization in order to maximize the margin width.

4.2. **Method.** We used scikit-learn's implementation of SVM. This choice was made due to scikit-learn's popularity and ease of use. We arbitrarily chose the rbf function as the nonlinear transformation K. [1]

Input Vectors. We arbitrarily chose 1000 as the largest integer, and so we only attempted to classify integers between 2 and 1000 (inclusive). Since SVM requires vector inputs, we will use the digits of the base-b representation of an integer as the entries of its input vector. Since SVM is scale invariant [3], we also scaled each input vector to $[-1, 1]$.

Example 4.2.1. Suppose that we are using base 3. Then $(5, 2)$ is the input vector for $32 \in \mathbb{Z}$. Likewise, the integers 153 and 1000 are represented by $(4, 1, 3)$ and $(4, 3, 4, 4)$, respectively. Since 1000 is the largest integer that we are interesting in, $(4, 3, 4, 4)$ is the longest vector in this input vector space.

[1]Scikit-learn has four built-in functions: linear, polynomial, rbf, and sigmoid. We believe that this classification problem requires a custom function. However, much more research is needed in order to find a function that can separate the primes from the composites.

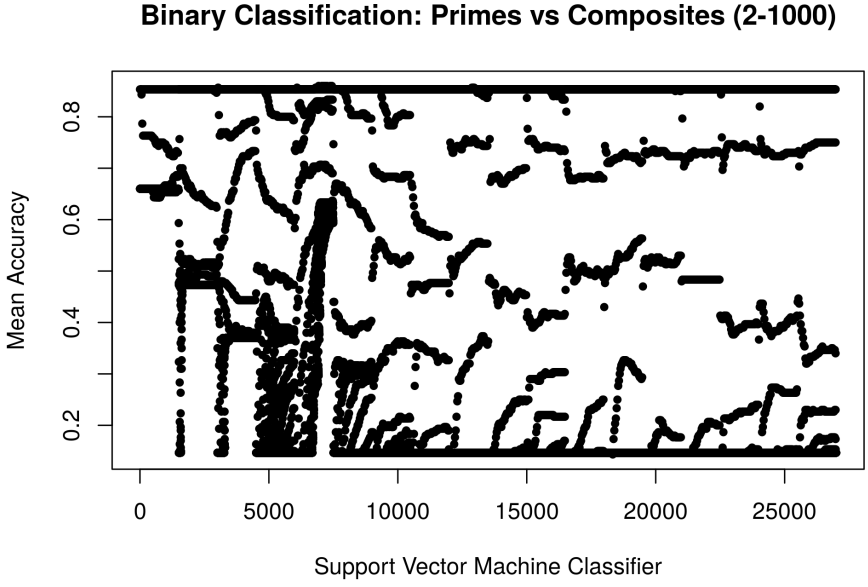


FIGURE 3. Complexity vs Mean Accuracy

Finding the Best Parameters. We used trial and error to find the base-b representation and SVM parameters that would maximize the mean accuracy of the algorithm. We experimented with base-b representations ranging from 2 to 500. We also considered different combinations of the following SVM parameters:

- class weight for input vectors that represent prime numbers, and
- penalty parameter C.

Usually, each class of input vectors has the same weight. However, there are only 168 primes between 2 and 1000. Thus, we tried different values for weight of the prime class. The penalty parameter C is the upper bound of the Lagrangian multipliers used in the optimization problem describe above. For each of these values, we tried values between 1 and 100.

Results. We refer to a particular combination of parameters as a model. The best models had a mean accuracy[2] equal to 0.86. Interestingly, all of these models used base 6 representations and a prime class weight of 2, with the penalty parameter C ranging from 3.2 to 4.9 (see appendix C).

A mean accuracy of 0.86 seems like a promising result. However, fig. 3 shows that a large number of models achieved a mean accuracy greater than 0.8 regardless of the parameters used (notice the thick line near the top). This lack of correlation is alarming. Since 82.8 percent of the integers between 2 and 1000 are composite, we suspect that the best models were simply classifying every integer as composite. Conversely, many models classified every integer as prime (notice the thick line near the bottom).

[2]Accuracy is the function $f : \mathbb{N}^2 \to \mathbb{R} \cap [0, 1]$ defined by $f(a, b) := \frac{a}{a+b}$, where a is the number of correct predictions and b is the number of incorrect predictions.

5. Conclusion

We conjecture that SVM's performance would improve and possibly exceed common tests for compositeness with the right transformation K. However, if it exists, finding such a function that maps integers into a space where composite numbers and primes are separable will require much more research. Nevertheless, it is a worthwhile endeavor for number theorists and cryptologists alike. Pierre de Fermat first stated Fermat's Little Theorem nearly 400 years ago in a letter dated October 18, 1640 [4]. Remarkably, it remains at the heart of a test for compositeness that still outperforms modern machine learning algorithms! For good reason, programming languages like Maple and Wolfram implement the Miller-Rabin test.

Appendix A. Miller-Rabin Test Results for 169

The Miller-Rabin test was performed with each nonzero $\bar{a} \in 169$, starting with $\bar{1}$ and ending with $\overline{168}$. Of these, 156 are Miller-Rabin witnesses to the compositeness of 169, namely,

2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 71, 72, 73, 74, 75, 76, 77, 78, 79, 81, 82, 83, 84, 85, 86, 87, 88, 90, 91, 92, 93, 94, 95, 96, 97, 98, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 148, 149, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167.

Thus, more than 9 out of 10 elements in $\mathbb{Z}_{169}$ are witnesses. This illustrates the effectiveness of the Miller-Rabin test.

Appendix B. Finding the Exponents

The most difficult step in the Miller-Rabin test for compositeness is finding the exponents. However, the following code snippet demonstrates how easily this can be done in Python.

```
# choose a and initialize n
a, n = 154320140719312, 1
# compute first exmponent
exp = a / pow(2, n)
# while 2 divides the expoenent
while exp % 1 == 0:
# print result
print(exp)
# increment n
n = n + 1
# compute next exponent
exp = a / pow(2, n)
```

Appendix C. Best Parameters

Figure 4. Mean accuracy converged to smallest value for bases greater than 15.

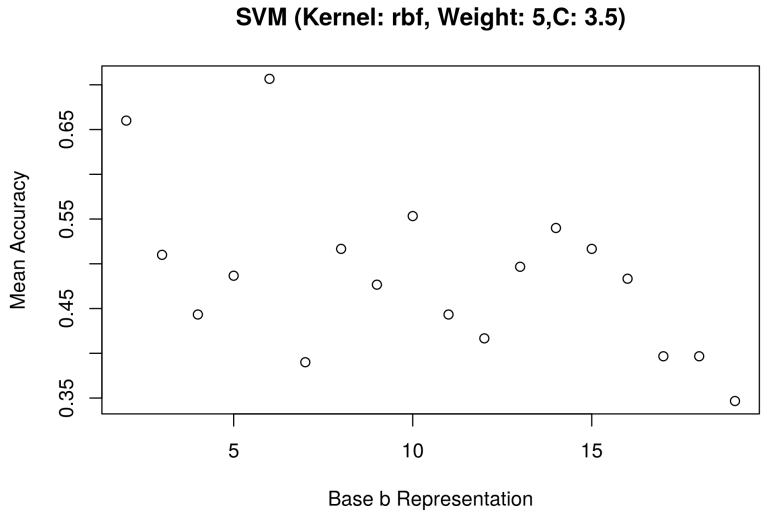


Figure 5. The best prime class weights were less than 3.

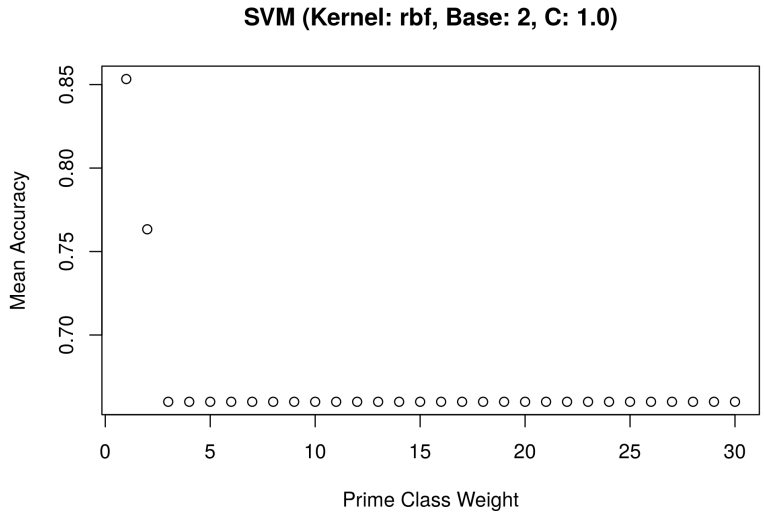
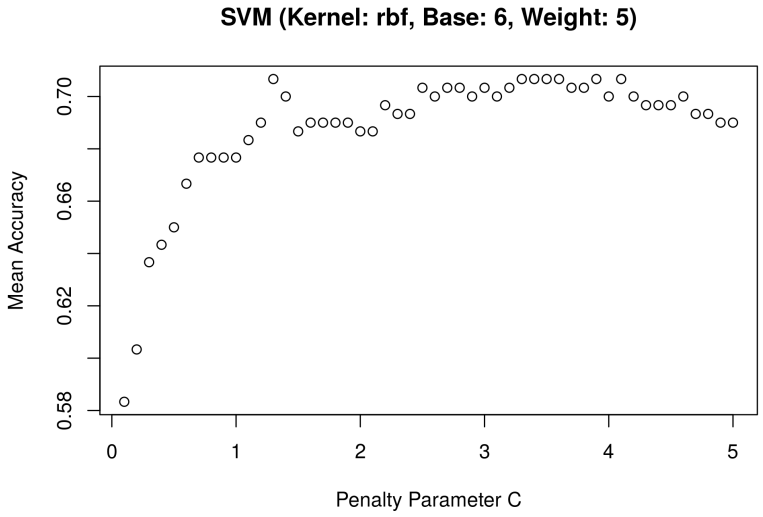


FIGURE 6. Models with penalty parameters C between 3 and 5 had the greatest mean accuracy.



LIST OF TABLES

References

[1] Carl Friedrich Gauss. Disquisitiones Arithmeticae. Yale University Press, 1966. Translated by Arthur A. Clarke, S.J.

[2] T. Marks J. Pommersheim and E. Flapan. Number Theory: A Lively Introduction with Proofs, Applications, and Stories. John Wiley & Sons, 2010.

[3] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12:2825–2830, 2011.

[4] Wikipedia. Fermat's little theorem — Wikipedia, The Free Encyclopedia, 2017. [Online; accessed 19-May-2017].

[5] Wikipedia. Support vector machine — Wikipedia, The Free Encyclopedia, 2017. [Online; accessed 19-May-2017].

[6] Patrick Winston. 6.034 Artificial Intelligence. Fall 2010. Massachusetts Institute of Technology: MIT OpenCourseWare, https://ocw.mit.edu. License: Creative Commons BY-NC-SA.