Analogies between the Real and Digital Lines and Circles

Dane M. Lawhorne

submitted in partial fulfillment of the requirements for Honors in Mathematics at the University of Mary Washington

Fredericksburg, Virginia

April 2014

This thesis by **Dane M. Lawhorne** is accepted in its present form as satisfying the thesis requirement for Honors in Mathematics.

Date	Approved
____________	______________________________ Randall D. Helmstutler, Ph.D. thesis advisor
____________	______________________________ Janusz Konieczny, Ph.D. committee member
____________	______________________________ J. Larry Lehman, Ph.D. committee member

Contents

1 Introduction **1**

2 Definitions and Basic Properties **1**
2.1 Definitions . . . 1
2.2 Separation Axioms and the Alexandrov Property . . . 3
2.3 Path Connectedness and Compactness . . . 4

3 Isometry and Automorphism Groups **6**
3.1 The Isometry Group of the Real Line . . . 6
3.2 The Automorphism Group of the Digital Line . . . 7
3.3 Semidirect Products and the Dihedralization . . . 10
3.4 The groups $\mathrm{Iso}(\mathbb{R})$ and $\mathrm{Aut}(\mathbb{D})$ as Dihedralizations . . . 10

4 Fundamental Groups **12**
4.1 The Construction of the Fundamental Group . . . 12
4.2 The Fundamental Groups of the Real Line and S^1 . . . 14
4.3 The Fundamental Groups of the Digital Line and Circles . . . 17

5 Conclusion **22**

References **24**

Abstract

The topological spaces known as the digital line and digital circles were created by mathematicians in the 1970's. Although these spaces are primarily studied for their applications, we adopt a theoretical point of view and establish several analogies between the digital line and circles and the real line and unit circle S^1. We first show that the automorphism group of the digital line has a structure analogous to that of the isometry group of the real line. Specifically, we prove that both of these groups are isomorphic to the dihedralization of their subgroup of translations. We then show that, like the real line, the digital line is simply connected. Next, we demonstrate that the digital line is a covering space of each digital circle by constructing a covering map which is analogous to the standard periodic covering map from the real line to S^1. Finally, we use this covering map to prove that the digital circles and S^1 have isomorphic fundamental groups.

1 Introduction

In the 1970's, mathematicians and computer scientists began using topological ideas to study digital images. To do this, they constructed spaces which functioned as topological models of digital image displays. One of the most important of these spaces is the *digital line*, which represents an infinite line of pixels. Other important digital spaces include the *digital circles*, which are quotient spaces of subsets of the digital line. An introduction to digital topology, as well as a brief history of the subject, may be found in [6].

In this paper, we will investigate which properties are lost and which are preserved when we transition from the real line to the digital line (and from the unit circle S^1 to the digital circles). As we will see, the digital line and circles have many algebraic and homotopical properties which are analogous to (algebraic and homotopical) properties of the real line and S^1. We will establish four of these analogies. First, both the isometry group of the real line and the automorphism group of the digital line are isomorphic to the dihedralization of a subgroup of functions called *translations*. Second, both the digital line and real line are simply connected. Third, there exists a covering map from the digital line to each digital circle which is analogous to the standard covering map from the real line to S^1 (both covering maps can be thought of as wrapping the covering space around the base space). Finally, we show that the digital circles and S^1 have isomorphic fundamental groups. These analogies provide evidence that the digital line and circles, at least from the viewpoint of basic algebraic topology, can be thought of as "pixelated" versions of the real line and S^1. Throughout this paper, we will refer to the real line and S^1 as the *Euclidean counterparts* of the digital line and circles.

2 Definitions and Basic Properties

We begin with definitions of the digital line and digital circles. We then examine a few important differences between these spaces and their Euclidean counterparts. Finally, we show that the digital line and circles behave in the same manner as their Euclidean counterparts with respect to path connectedness and compactness.

2.1 Definitions

Let X be a topological space. Recall that a *basis* $\mathcal{B}$ for the topology on X is a collection of open sets such that any open set in X can be written as a union of sets from $\mathcal{B}$. Elements of $\mathcal{B}$ are called *basis elements*. We define the digital line by specifying its basis elements.

Definition 2.1. The *digital line* is the set of integers with basis elements $B(n)$ defined by

$$B(n) = \begin{cases} \{n\} & \text{if } n \text{ is odd} \\ \{n-1, n, n+1\} & \text{if } n \text{ is even} \end{cases}$$

for each integer n. We use $\mathbb{D}$ to denote the digital line.

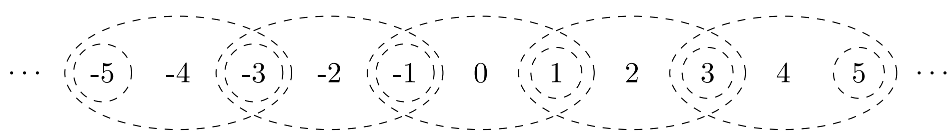
Every open set in the digital line is a union of basis elements $B(n)$. Other authors use $\mathbb{Z}$ to denote the digital line, but we refrain from using this notation to avoid confusion with the additive group of integers.

We now verify that the basis elements from Definition 2.1 do in fact generate a valid topology on the integers. Recall that a basis for a topology on a set X must satisfy the following two conditions:

1. Every $x \in X$ is contained in a basis element.
2. If x is contained in an intersection of basis elements $B_1 \cap B_2$, then there exists a basis element B_3 such that $x \in B_3$ and $B_3 \subseteq B_1 \cap B_2$.

If both of these conditions are satisfied, then the basis elements generate a valid topology. It is clear that any integer n is contained in a basis element from Definition 2.1; specifically, n is contained in $B(n)$. Now suppose $B(n) \cap B(m)$ is a nonempty intersection. Clearly n and m cannot both be odd since $B(n) \cap B(m)$ is nonempty. If one of n and m, say m, is odd and the other is even, then $m = n + 1$ or $m = n - 1$, and $B(n) \cap B(m) = \{m\} = B(m)$. If both n and m are even, then m and n are consecutive even integers, and $B(n) \cap B(m) = \{n + 1\} = B(n + 1)$ if $n < m$ and $B(n) \cap B(m) = \{m + 1\} = B(m + 1)$ if $n > m$. Hence the basis elements from Definition 2.1 generate a valid topology on the integers.

We can visualize the digital line by encircling its basis elements:



The digital line

As we mentioned in our introduction, the digital line represents an infinite line of pixels. Specifically, odd integers represent the pixels themselves and even integers represent the boundaries between pixels. Further information about this definition may be found in [1, p. 45]. Note that all sets $\{n\}$ where n is even are closed sets in the digital line.

The next spaces we will study are the digital circles, which are formed from subsets of the digital line called *digital intervals*.

Definition 2.2. The *digital interval* $[a, b]$ is the set $\{n \in \mathbb{D} \mid a \leq n \leq b\}$.

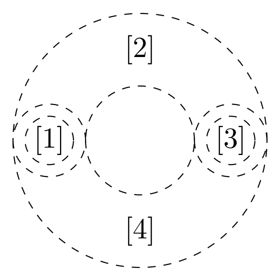
Note that the digital interval $[a, b]$ is open in the digital line if both a and b are odd and closed if both a and b are even. If one of a and b is odd and the other is even, then $[a, b]$ is neither open nor closed.

Definition 2.3. Given an even integer $n \geq 4$, the *digital circle* C_n is the quotient space of the digital interval $[1, n + 1]$ (with the subspace topology) under the single identification $1 \sim n + 1$.

Each digital circle C_n has n points, which are equivalence classes, and except for the class $[1] = [n+1]$, each class has a unique representative integer. Furthermore, if $[a] \neq [n+1]$, then $\{[a]\}$ is open if and only if $q^{-1}([a]) = \{a\}$ is open, where $q : [1, n+1] \to C_n$ is the quotient map. This means that $\{[a]\}$ is open if and only if a is odd. The set $\{[n+1]\}$ is also open since $q^{-1}(\{[n+1]\}) = \{1, n+1\} = \{1\} \cup \{n+1\}$ is open (since n is even). Additionally, if a is even, any open set U containing $[a]$ must also contain $[a-1]$ and $[a+1]$ since $q^{-1}(U)$ is an open set containing a, and thus contains $a-1$ and $a+1$. We can now define the smallest open set $B([a])$ to which class $[a]$ belongs. If a is odd, then $B([a]) = \{[a]\}$. If a is even, then $B([a]) = \{[a-1], [a], [a+1]\}$. Furthermore, it is easy to see that the set $\{B([a]) \mid [a] \in C_n\}$ is a basis for the topology on C_n.

The construction of the digital circles is analogous to the construction of S^1 as a quotient space of the closed (real) interval $[0, 1]$ under the single identification $0 \sim 1$. We can make this analogy even stronger by constructing each digital circle as the quotient space of the *closed* digital interval $[2, n+2]$ under the single identification $2 \sim n+2$. The resulting space D_n is homeomorphic to C_n via the homeomorphism $h : D_n \to C_n$ defined by $h([a]) = [b-2]$ where b is the largest element of $[a]$. That is, if $[a] \neq [2] = [n+2]$, $h([a]) = [a-2]$, and $h([2]) = h([n+2]) = [n]$.

The smallest digital circle is the four-point circle C_4, which is the quotient space of the digital interval $[1, 5]$ under the single identification $1 \sim 5$. The open sets in C_4 are $\varnothing$, $\{[1]\}$, $\{[3]\}$, $\{[1], [2], [3]\}$, $\{[1], [3], [4]\}$ and C_4 itself. We can visualize this space by encircling its basis elements:



The four-point digital circle C_4

It is clear from this visualization that the topology on C_4 is locally similar to the digital line's topology. We will see that this similarity holds for each digital circle, which allows us to prove the existence of an important relationship between the digital line and circles, namely that the digital line is a *covering space* of each digital circle.

2.2 Separation Axioms and the Alexandrov Property

Before we examine algebraic and homotopical similarities between the digital line and circles and their Euclidean counterparts, we first discuss a few important differences. The first main difference is that unlike the real line and S^1, the digital line and circles are not Hausdorff. In the digital line, any open set containing an even integer n also contains $n+1$. So if an even integer n is contained in a open set U and $n+1$ is contained in an open set V, it follows that $n+1 \in U \cap V$. Hence the digital line is not Hausdorff. Similarly, any open set containing $[a] \in C_n$ where a is even also contains $[a+1]$, so a similar argument shows that each digital circle is not Hausdorff.

Fortunately, the digital line and circles do satisfy a weaker separation axiom, called the T_0 *property*.

Definition 2.4. A space X is T_0 if, for any two points in X, there exists an open set containing one of the points which does not contain the other.

Proposition 2.5. *The digital line is* T_0*.*

Proof. Let $n, m \in \mathbb{D}$ with $n \neq m$. Suppose either n or m is odd. Say without loss of generality that n is odd. Then the basis element $B(n)$ contains n and does not contain m since it is a one-point set. If both n and m are even, then the open set $\{n-1, n, n+1\}$ does not contain m. It follows that the digital line is T_0. □

Proposition 2.6. *Each digital circle is T_0.*

The proof of Proposition 2.6 is similar to the proof of Proposition 2.5.

There is another separation axiom which is stronger than the T_0 property but weaker than the Hausdorff property. Recall that a space X is T_1 if for any $x, y \in X$, there exists an open set U such that $x \in U$ and $y \notin U$ and an open set V such that $y \in V$ and $x \notin V$. Our argument that the digital line and circles are not Hausdorff also shows that they are not T_1.

At this point, we see that the real line and S^1 have nice properties (the Hausdorff and T_1 properties) which the digital line and circles do not. However, there is another useful property, called the *Alexandrov property*, which the digital line and circles have but their Euclidean counterparts do not.

Definition 2.7. A space X is *Alexandrov* if arbitrary intersections of open sets are open.

The intersection $\bigcap_{n=1}^{\infty}\left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$ shows that the real line is not Alexandrov. A similar approach would show that S^1 is not Alexandrov. For example, we could take the intersection of infinitely many shrinking open arcs about the point $(0,1)$. However, each digital circle is clearly Alexandrov since it is finite (the only possible intersections of open sets are finite intersections).

To see that that the digital line is Alexandrov, it will be helpful to introduce an equivalent definition of the Alexandrov property. We say that a space X has the *minimal open set property* if for each point $x \in X$, there exists an open set M_x containing x such that $M_x \subseteq U$ for any other open set U containing x. If X has the minimal open set property, it is not difficult to see that an arbitrary intersection of open sets in X must be equal to the union of the minimal open sets M_x for each x contained in that intersection, and is therefore open. Furthermore, if X is Alexandrov, then the intersection of all open sets containing $x \in X$ is the minimal open set about x. Hence the minimal open set property is equivalent to the Alexandrov property.

Proposition 2.8. *The digital line is Alexandrov.*

Proof. The digital line clearly has the minimal open set property: for any $n \in \mathbb{D}$, the basic open set $B(n)$ is the minimal open set containing n. Hence the digital line is Alexandrov. □

2.3 Path Connectedness and Compactness

The digital line and circles behave in the same manner as their Euclidean counterparts with respect to path connectedness and compactness. Recall that both the real line and S^1 are path connected, S^1 is compact, and the real line is not compact.

Proposition 2.9. *The digital line is path connected.*

Proof. Let $n \in \mathbb{D}$. If n is odd, define a function $\beta : [0,1] \to \mathbb{D}$ by

$$\beta(t) = \begin{cases} n & \text{if } t \in [0,1) \\ n+1 & \text{if } t = 1. \end{cases}$$

If β is continuous, then it is a path from n to $n+1$ since $\beta(0) = n$ and $\beta(1) = n+1$. To verify that β is continuous, it is sufficient to check that inverse images of basis elements are open. The only basis elements which overlap with the image of β are $B(n-1) = \{n-2, n-1, n\}$, $B(n) = \{n\}$, and $B(n+1) = \{n, n+1, n+2\}$. Since $\beta^{-1}(\{n-2, n-1, n\}) = [0,1)$ is open, $\beta^{-1}(\{n\}) = [0,1)$ is open, and $\beta^{-1}(\{n, n+1, n+2\}) = [0,1]$ is open, β is indeed continuous. Hence β is a path from n to $n+1$.

If n is even, define a function $\gamma : [0,1] \to \mathbb{D}$ by

$$\gamma(t) = \begin{cases} n & \text{if } t = 0 \\ n+1 & \text{if } t \in (0,1]. \end{cases}$$

If γ is continuous, then it is a path from n to $n+1$ since $\gamma(0) = n$ and $\gamma(1) = n+1$. Again, it is sufficient to check that inverse images of basis elements are open, so the only sets we need to check are $B(n) = \{n-1, n, n+1\}$ and $B(n+1) = \{n+1\}$. Since $\gamma^{-1}(\{n-1, n, n+1\}) = [0,1]$ and $\gamma^{-1}(\{n+1\}) = (0,1]$, γ is indeed continuous.

Recall that in any space X, we can define an equivalence relation $\sim$ by saying $x \sim y$ if there exists a path in X from x to y. In the digital line, we have shown that $n \sim n+1$ for any n. For any two distinct points $n_1, n_2 \in \mathbb{D}$ (supposing without loss of generality that $n_2 > n_1$), $n_2 = n_1 + k$ for some positive integer k. Since $n_1 \sim n_1 + 1 \sim n_1 + 2 \sim \cdots \sim n_1 + k$, it follows that $n_1 \sim n_2$. Hence there exists a path between any two points in the digital line, and so the digital line is path connected. □

A similar argument shows that the digital interval $[1, n+1]$ is path connected for any n. Since quotients preserve path connectedness and each digital circle is a quotient space of $[1, n+1]$ for some even integer $n \geq 4$, each digital circle is also path connected.

Our arguments regarding compactness are more straightforward. The collection of digital intervals $\{[-n, n] \mid n \text{ is odd}\}$ is clearly an irreducible open cover of the digital line. Thus we have:

Proposition 2.10. *The digital line is not compact.*

Finally, each digital circle is clearly compact since it is finite. We can also characterize which subsets of the digital line are compact and which are connected.

Proposition 2.11. *A subset of the digital line is connected if and only if it is a set of consecutive integers.*

Proof. Let A be a connected subset of the digital line. If A were not a set of consecutive integers, there would exist some $c \notin A$ such that A would contain numbers less than c and greater than c. But then

$$\{x \mid x \in A, x < c\} \cup \{x \mid x \in A, x > c\}$$

would be a separation of A, contradicting connectedness. Hence A is a set of consecutive integers. The proof of the reverse direction is similar to the proof of Proposition 2.9. □

Proposition 2.12. *A subset of the digital line is compact if and only if it is finite.*

Proof. Let A be a subset of the digital line. If A were not finite,

$$\bigcup_{n \in A} B(n) \cap A$$

would be an irreducible open cover of A, contradicting compactness. Hence A is a finite set of consecutive integers. The proof of the reverse direction is obvious. □

3 Isometry and Automorphism Groups

An important group associated with a topological space X is the *automorphism group* $\mathrm{Aut}(X)$, which is the set of all self-homeomorphisms $f : X \to X$ under the operation of function composition. If X is a metric space, we can also construct the *isometry group* $\mathrm{Iso}(X)$, which consists of all distance-preserving surjections $f : X \to X$ under function composition. If X is a metric space, then $\mathrm{Iso}(X)$ is a subgroup of $\mathrm{Aut}(X)$.

In this section, we examine similarities between the isometry group of the real line and the automorphism group of the digital line. Both groups consist entirely of functions called *translations* and *reflections*. Furthermore, both $\mathrm{Iso}(\mathbb{R})$ and $\mathrm{Aut}(\mathbb{D})$ are isomorphic to the *dihedralizations* of their subgroup of translations. The dihedralization generalizes the structure of the familiar dihedral groups (the symmetry groups of the regular n-gons).

3.1 The Isometry Group of the Real Line

We begin with definitions of an *isometry* and the *isometry group*.

Definition 3.1. Let (X, d_X) and (Y, d_Y) be metric spaces. An *isometry* from X to Y is a function $f : X \to Y$ such that $d_Y(f(a), f(b)) = d_X(a, b)$ for all $a, b \in X$.

Isometries are always injective but not necessarily surjective. To form the isometry group, we require that isometries be surjective, and hence bijective, so that each element has an inverse. With this additional requirement, it is straightforward that the set of all isometries from a metric space X to itself satisfies the axioms for a group. The composition of two surjective isometries is a surjective isometry and composition of functions is always associative. The identity element is the identity map 1_X, and the inverse of a surjective isometry f is f^{-1}.

Definition 3.2. Let X be a metric space. The *isometry group* of X, denoted $\mathrm{Iso}(X)$, is the set of all surjective isometries $f : X \to X$ under function composition.

We are interested in surjective isometries from the real line to itself. Some surjective isometries are known as *translations*, which can be visualized as shifts of the real line by a specified real number (for us, positive numbers indicate a shift to the right and negative numbers indicate a shift to the left). Other surjective isometries are known as *reflections*, which can be visualized by reflecting the real line 180 degrees about a specified real number.

Definition 3.3. The *translation* of the real line by the real number a is the function $T_a : \mathbb{R} \to \mathbb{R}$ defined by $T_a(x) = x + a$. The *reflection* of the real line about the real number a is the function $R_a : \mathbb{R} \to \mathbb{R}$ defined by $R_a(x) = 2a - x$.

There are two important facts about translations and reflections which will be used in later computations. The first is that $(T_a)^{-1} = T_{-a}$ for any translation T_a. The second is that $(R_a)^{-1} = R_a$ for any reflection R_a. It follows that the subgroup $\langle R_a \rangle$ of $\mathrm{Iso}(\mathbb{R})$ generated by a reflection is a group of order two.

It turns out that every surjective isometry $f : \mathbb{R} \to \mathbb{R}$ can be given one of these geometric interpretations. To prove this, we will use a special type of isometry called a *linear isometry*, which is an isometry $f : \mathbb{R} \to \mathbb{R}$ such that $f(0) = 0$. Note that if $g : \mathbb{R} \to \mathbb{R}$ is a surjective isometry (not necessarily linear), the function $g_0 : \mathbb{R} \to \mathbb{R}$ defined by $g_0(x) = g(x) - g(0)$ is a surjective linear isometry. Furthermore, g can be written as the composition $g = T_{g(0)} \circ g_0$. That is, any surjective isometry from the real line to itself is the composition of a translation and a linear isometry. A few simple calculations give us the equations $T_a \circ T_b = T_{a+b}$ and $T_a \circ R_b = R_{\frac{1}{2}a+b}$ for any real numbers

a and b. So if every linear isometry is a translation or a reflection, then every surjective isometry is a translation or a reflection. We now show that a linear isometry is indeed a translation or a reflection. In fact, any linear isometry is either the translation T_0 (which is the identity map $1_{\mathbb{R}}$) or the reflection R_0 (which maps x to $-x$).

Proposition 3.4. *Let $f : \mathbb{R} \to \mathbb{R}$ be a linear isometry. Then $f = T_0$ or $f = R_0$.*

Proof. Let $f : \mathbb{R} \to \mathbb{R}$ be a linear isometry. For all $x \in \mathbb{R}$, $|f(x) - f(0)| = |x - 0|$, and so $|f(x)| = |x|$. For each $x \in \mathbb{R}$, either $f(x) = x$ or $f(x) = -x$. Specifically, $f(1) = 1$ or $f(1) = -1$.

Suppose first that $f(1) = 1$. We will show that $f(x) = x = T_0(x)$ for all $x \in \mathbb{R}$. We already know that $f(x) = x$ or $f(x) = -x$. Suppose there exists some $x \in \mathbb{R}$ such that $f(x) = -x$. Then $|f(1) - f(x)| = |1 - (-x)| = |1 + x|$, and so $|1 + x| = |1 - x|$. But this implies $x = 0$. In other words, if $f(x) = -x$, then $x = 0$. It follows that $f(x) = x = T_0(x)$ for all $x \in \mathbb{R}$, and so $f = T_0$.

Now suppose that $f(1) = -1$. We will show that $f(x) = -x = R_0(x)$ for all $x \in \mathbb{R}$. Again, we already know that $f(x) = x$ or $f(x) = -x$. Suppose there exists some $x \in \mathbb{R}$ such that $f(x) = x$. Then $|f(1) - f(x)| = |-1 - x|$, and so $|-1 - x| = |1 - x|$. Again, this implies $x = 0$. So the only real number x mapped by f to itself is 0. It follows that $f(x) = -x = R_0(x)$ for all $x \in \mathbb{R}$, and so $f = R_0$. □

For reasons already stated, we now have:

Proposition 3.5. *Any surjective isometry $f : \mathbb{R} \to \mathbb{R}$ is a translation or a reflection.*

Let $T(\mathbb{R})$ denote the set of all translations of the real line. We now show that $\mathrm{Iso}(\mathbb{R})$ is generated by $T(\mathbb{R})$ and the reflection R_0.

Proposition 3.6. *The isometry group of the real line is generated by the set $T(\mathbb{R})$ of translations and the reflection R_0.*

Proof. It is suffient to prove that any reflection is the composition of R_0 and a translation. Let $R_a : \mathbb{R} \to \mathbb{R}$ be a reflection. For every $x \in \mathbb{R}$, we have

$$\begin{aligned} R_a(x) &= 2a - x \\ &= -(x - 2a) \\ &= R_0(x - 2a) \\ &= R_0(T_{-2a}(x)) \\ &= (R_0 \circ T_{-2a})(x) \end{aligned}$$

so $R_a = R_0 \circ T_{-2a}$. Hence $T(\mathbb{R})$ and R_0 generate $\mathrm{Iso}(\mathbb{R})$. □

3.2 The Automorphism Group of the Digital Line

Having sufficiently described the isometry group of the real line in terms of its generators, we now focus our attention on the automorphism group of the digital line. We will see that the structure of $\mathrm{Aut}(\mathbb{D})$ is similar to the structure of $\mathrm{Iso}(\mathbb{R})$. We begin with definitions of an *automorphism* and the *automorphism group.*

Definition 3.7. An *automorphism* of a topological space X is a homeomorphism from X to itself.

It is straightforward that the set of all automorphisms of X satisfy the axioms for a group. The composition of two automorphisms is itself an automorphism, and composition of functions is always associative. The identity element is the identity map 1_X, and the inverse of an automorphism f is f^{-1}.

Definition 3.8. The *automorphism group* of a space X, denoted $\text{Aut}(X)$, is the set of all automorphisms $f : X \to X$ under function composition.

We now define translations and reflections of the digital line, which have the same geometric interpretation as translations and reflections of the real line.

Definition 3.9. The *translation* of the digital line by the even integer a is the function $\tau_a : \mathbb{D} \to \mathbb{D}$ defined by $\tau_a(x) = x + a$. The *reflection* of the digital line about the (even or odd) integer a is the function $\rho_a : \mathbb{D} \to \mathbb{D}$ defined by $\rho_a(x) = 2a - x$.

Recall that translations of the real line $T_a : \mathbb{R} \to \mathbb{R}$ are defined for any real number a. In the the digital line, we can only translate by even integers because tranlsations by odd integers are not continuous. The loss of continuity results from the fact that odd integers (which are open by themselves) pull back to even integers (which are not open by themselves). However, we can reflect the digital line about both odd and even integers without losing continuity.

Note that $(\tau_a)^{-1} = \tau_{-a}$ for any translation τ_a and that $(\rho_a)^{-1} = \rho_a$ for any reflection ρ_a. It follows that the subgroup $\langle \rho_a \rangle$ of $\text{Aut}(\mathbb{D})$ generated by a reflection is a group of order two.

Just like $\text{Iso}(\mathbb{R})$, we will see that $\text{Aut}(\mathbb{D})$ consists entirely of translations and reflections. The proof, however, uses strong induction and requires slightly more tedious calculations. Before giving this proof, we first show that any continuous bijection $f : \mathbb{D} \to \mathbb{D}$ is *parity-preserving*, which means that odd points are mapped to odd points and even points are mapped to even points.

Lemma 3.10. *Any continuous bijection from the digital line to itself is parity-preserving.*

Proof. Let $f : \mathbb{D} \to \mathbb{D}$ be a continuous bijection. If n is odd, then $\{n\}$ is open, as is $f^{-1}(\{n\})$. Since f is bijective, $f^{-1}(\{n\})$ is a one-point set. Since the only open one-point sets are sets containing an odd point, $f^{-1}(n)$ is odd. This means that even points cannot map to odd points, and therefore map to even points. Similarly, if m is even, then $\{m\}$ is closed, as is $f^{-1}(\{m\})$. Since f is bijective, $f^{-1}(\{m\})$ is a one-point set. Since the only closed one-point sets are sets containing an even point, $f^{-1}(m)$ is even. Hence odd points cannot map to even points, and therefore map to odd points. It follows that f is parity-preserving. □

Proposition 3.11. *Any automorphism of the digital line is a translation or a reflection.*

Proof. Let $f : \mathbb{D} \to \mathbb{D}$ be an automorphism. Let $n = f(0)$. Since f is continuous, we know that n is even, and so $\{n-1, n, n+1\}$ is open. This means that $f^{-1}(\{n-1, n, n+1\})$ is also open. Since $0 \in f^{-1}(\{n-1, n, n+1\})$ and f is bijective, it follows from the digital line's topology that $f^{-1}(\{n-1, n, n+1\}) = \{-1, 0, 1\}$. We now have two cases. The first case is that $f(1) = n+1$ and $f(-1) = n-1$, and the second is that $f(1) = n-1$ and $f(-1) = n+1$.

Suppose for the first case that $f(1) = n+1$ and $f(-1) = n-1$. We will show that $f = \tau_n$. Suppose, using strong induction, that for all $0 \le w \le x$, $f(w) = n + w$ (note that the base case is valid since $f(1) = n+1$). We will show that $f(x+1) = n+x+1$. This case breaks down into two subcases, depending on the parity of x.

Suppose first that x is odd. Then $n+x$ is also odd since n is even. So $\{n+x, n+x+1, n+x+2\}$ is open, as is $f^{-1}(\{n+x, n+x+1, n+x+2\})$. Since $f(x) = n+x$, $f(x-1) = n+x-1$, and f is bijective, $f^{-1}(\{n+x, n+x+1, n+x+2\})$ is a three-point open set containing x which does not contain $x-1$. Furthermore, $f^{-1}(\{n+x, n+x+1, n+x+2\})$ also contains an even integer since $n+x+1$ is even and f is parity-preserving. It follows that $f^{-1}(\{n+x, n+x+1, n+x+2\}) = \{x, x+1, x+2\}$. Since f is parity-preserving, it follows that $f(x+1) = n+x+1$, which is our desired result.

Suppose now that x is even. Then $n+x$ is also even since n is even. So $\{n+x-1, n+x, n+x+1\}$ is open, as is $f^{-1}(\{n+x-1, n+x, n+x+1\})$. From our assumption, we know that $f(x-1) = n+x-1$

and $f(x) = n + x$. Since f is bijective, $f^{-1}(\{n+x-1, n+x, n+x+1\})$ is a three-point open set containing $x-1$ and x. It follows that $f^{-1}(\{n+x-1, n+x, n+x+1\}) = \{x-1, x, x+1\}$, and so $f(x+1) = n+x+1$. Using strong induction, it follows for both subcases that $f(x) = n+x = \tau_n(x)$ for any positive integer x. Using a similar argument, we could also conclude that $f(x) = n+x = \tau_n(x)$ for any negative integer x. Hence $f = \tau_n$.

Suppose for the second case that $f(1) = n-1$ and $f(-1) = n+1$. We will show that $f = \rho_{\frac{n}{2}}$. Suppose, using strong induction, that for all $0 \leq w \leq x$, $f(w) = n - w$ (note again that the base case is valid). We will show that $f(x+1) = n - (x+1)$. Again, this case breaks down into two subcases, depending on whether x is odd or even.

Suppose first that x is odd. Then $n - x$ is also odd. So $\{n-x-2, n-x-1, n-x\}$ is open, as is $f^{-1}(\{n-x-2, n-x-1, n-x\})$. We know from our assumption that $f(x) = n - x$ and $f(x-1) = n-x+1$. Since f is bijective, $f^{-1}(\{n-x-2, n-x-1, n-x\})$ is a three-point open set containing x which does not contain $x-1$. Furthermore, $f^{-1}(\{n-x-2, n-x-1, n-x\})$ contains an even integer since f is parity-preserving. It follows that $f^{-1}(\{n-x-2, n-x-1, n-x\}) = \{x, x+1, x+2\}$. Since f is parity-preserving, we have $f(x+1) = n-x-1 = n-(x+1)$, which is our desired result.

Suppose now that x is even. Then $n-x$ is also even. So $\{n-x-1, n-x, n-x+1\}$ is open, as is $f^{-1}(\{n-x-1, n-x, n-x+1\})$. We know that $f(x) = n-x$ and $f(x-1) = n-x+1$, and so $f^{-1}(\{n-x-1, n-x, n-x+1\})$ is a three-point open set containing x and $x-1$. It follows that $f^{-1}(\{n-x-1, n-x, n-x+1\}) = \{x-1, x, x+1\}$. Since f is bijective, $f(x+1) = n-x-1 = n-(x+1)$. Using strong induction, it follows for both subcases that $f(x) = n - x = \rho_{\frac{n}{2}}(x)$ for any positive integer x. A similar argument would show that $f(x) = n - x = \rho_{\frac{n}{2}}(x)$ for any negative integer x. Hence, $f = \rho_{\frac{n}{2}}$.

We have now shown that for any automorphism $f : \mathbb{D} \to \mathbb{D}$, $f = \tau_n$ or $f = \rho_{\frac{n}{2}}$, where $n = f(0)$. In other words, f is either a translation or a reflection of the digital line. $\square$

Recall that $\mathrm{Iso}(\mathbb{R})$ is generated by the set $T(\mathbb{R})$ of all translations of the real line and the reflection R_0. The automorphism group of the digital line is also generated by translations and the reflection ρ_0. In fact, $\mathrm{Aut}(\mathbb{D})$ has two generators: the translation τ_2 and the reflection ρ_0.

Proposition 3.12. *The automorphism group of the digital line is generated by τ_2 and ρ_0.*

Proof. Let $f : \mathbb{D} \to \mathbb{D}$ be an automorphism. Then $f = \tau_n$ for some even n or $f = \rho_n$ for some (even or odd) n. Suppose first that $f = \tau_n$ for some even n. Since n is even, $n = 2k$ for some integer k. Thus we have

$$\begin{aligned} f(x) &= x + n \\ &= x + 2k \\ &= \tau_{2k}(x) \\ &= (\tau_2)^k(x) \end{aligned}$$

so $f = (\tau_2)^k$. Now suppose that $f = \rho_n$ for some (even or odd) n. We have

$$\begin{aligned} f(x) &= 2n - x \\ &= -(x - 2n) \\ &= \rho_0(x - 2n) \\ &= \rho_0(\tau_{-2n}(x)) \\ &= (\rho_0 \circ \tau_{-2n})(x) \end{aligned}$$

so $f = \rho_0 \circ \tau_{-2n}$. Hence any automorphism of the digital line can be written as a composition of ρ_0 and τ_2. □

3.3 Semidirect Products and the Dihedralization

We have now fully described $\text{Iso}(\mathbb{R})$ and $\text{Aut}(\mathbb{D})$ in terms of their generators. It turns out that these groups can also be described by a construction called the *dihedralization.* A complete exposition of the dihedralization, along with motivation and examples, may be found in [2]. Before we introduce the dihedralization, we must describe the algebraic structures known as *actions* and *semidirect products.*

Let H and N be groups. An *action* of H on N is a homomorphism $\phi : H \to \text{Aut}(N)$, where $\text{Aut}(N)$ is the group consisting of all isomorphisms $f : N \to N$ under the operation of function composition (this group is called the *automorphism group* of N). Given an action $\phi : H \to \text{Aut}(N)$, we can define a new group $H \ltimes_\phi N$ as the set $H \times N$ under the following operation:

$$(h_1, n_1)(h_2, n_2) = (h_1 h_2, \phi_{h_2}^{-1}(n_1) n_2).$$

where ϕ_h is a shorthand notation for $\phi(h)$. The group $H \ltimes_\phi N$ is called the *external semidirect product* of H with N under ϕ.

There is also an internal version of this construction. If H and N are subgroups of a group G, let HN denote the set of all products hn where $h \in H$ and $n \in N$. If $G = HN$, $H \cap N = \{e\}$ (where e denotes the identity element of G), and N is normal in G, we say that G is the *internal semidirect product* of H and N and write $G = H \ltimes N$.

It turns out that any internal semi-direct product can be viewed as an external semi-direct product. Suppose $G = H \ltimes N$. Let $\phi : H \to \text{Aut}(N)$ be the action where $\phi_h : N \to N$ is defined by $\phi_h(n) = hnh^{-1}$ (we call this the *conjugation action*). It can be shown that $H \ltimes N \cong H \ltimes_\phi N$ via the isomorphism $f : H \ltimes N \to H \ltimes_\phi N$ defined by $f(hn) = (h, n)$. That is, every internal semidirect product can be viewed as an external semidirect product under the conjugation action. More in-depth treatments of semidirect products may be found in [5] and [3].

We now introduce the dihedralization of an abelian group.

Definition 3.13. Let G be an abelian group, and let C_2 denote the multiplicative group $\{\pm 1\}$. Let $\phi : C_2 \to \text{Aut}(G)$ be the action where $\phi_1 : G \to G$ is defined by $\phi_1(g) = g$ and and $\phi_{-1} : G \to G$ is defined by $\phi_{-1}(g) = g^{-1}$. The *dihedralization* of G, denoted $D(G)$, is the semidirect product $C_2 \ltimes_\phi G$.

We can only dihedralize abelian groups because the function $\phi_{-1} : G \to G$ defined by $\phi_{-1}(g) = g^{-1}$ is an automorphism if and only if G is abelian.

3.4 The groups $\text{Iso}(\mathbb{R})$ and $\text{Aut}(\mathbb{D})$ as Dihedralizations

We now show that the isometry group of the real line is isomorphic to the dihedralization $D(T(\mathbb{R}))$, where $T(\mathbb{R})$ denotes the subgroup of translations. We first show that $T(\mathbb{R})$ is indeed a subgroup, and in addition, that it is normal in $\text{Iso}(\mathbb{R})$. We then show that $\text{Iso}(\mathbb{R})$ is an internal semidirect product. Finally, we show that the externalization of this semidirect product is isomorphic to the dihedralization of $T(\mathbb{R})$.

Lemma 3.14. *The set $T(\mathbb{R})$ of all translations of the real line is a subgroup of* $\text{Iso}(\mathbb{R})$.

Proof. Let $T_a : \mathbb{R} \to \mathbb{R}$ and $T_b : \mathbb{R} \to \mathbb{R}$ be translations. Then $T_a \circ T_b = T_{a+b}$ is also a translation. Furthermore, $(T_a)^{-1} = T_{-a}$ is also a translation. Hence $T(\mathbb{R})$ is a subgroup of $\text{Iso}(\mathbb{R})$. □

Note that the subgroup $T(\mathbb{R})$ is also abelian since $T_a \circ T_b = T_{a+b} = T_b \circ T_a$.

Lemma 3.15. *The subgroup $T(\mathbb{R})$ of translations is normal in* $\mathrm{Iso}(\mathbb{R})$.

Proof. From the proof of Proposition 3.6, it is clear that $T(\mathbb{R})$ has index two in $\mathrm{Iso}(\mathbb{R})$. Since subgroups of index two are normal, $T(\mathbb{R})$ is normal in $\mathrm{Iso}(\mathbb{R})$. □

Lemma 3.16. *The isometry group of the real line line admits an internal semidirect product decomposition* $\mathrm{Iso}(\mathbb{R}) = \langle R_0 \rangle \ltimes T(\mathbb{R})$.

Proof. Let $f \in \mathrm{Iso}(\mathbb{R})$. Then f is either a translation T_a or a reflection R_a. If $f = T_a$, then $f = 1_{\mathbb{R}} \circ T_a$. If $f = R_a$, then $f = R_0 \circ T_{-2a}$ by Proposition 3.6. It follows that $\mathrm{Iso}(\mathbb{R}) = \langle R_0 \rangle T(\mathbb{R})$. Furthermore, $T(\mathbb{R})$ is normal in $\mathrm{Iso}(\mathbb{R})$ and $\langle R_0 \rangle \cap T(\mathbb{R}) = \{1_{\mathbb{R}}\}$. Hence $\mathrm{Iso}(\mathbb{R}) = \langle R_0 \rangle \ltimes T(\mathbb{R})$. □

Theorem 3.17. *The isometry group of the real line is isomorphic to the dihedralization $D(T(\mathbb{R}))$.*

Proof. We have already shown that $\mathrm{Iso}(\mathbb{R}) \cong \langle R_0 \rangle \ltimes T(\mathbb{R})$. This internal semidirect product may be viewed as the external semidirect product $\langle R_0 \rangle \ltimes_\phi T(\mathbb{R})$ where $\phi_{R_0} : T(\mathbb{R}) \to T(\mathbb{R})$ is defined by $\phi_{R_0}(T_a) = R_0 \circ T_a \circ (R_0)^{-1} = (T_a)^{-1}$ and $\phi_{1_{\mathbb{R}}} : T(\mathbb{R}) \to T(\mathbb{R})$ is defined by $\phi_{1_{\mathbb{R}}}(T_a) = 1_{\mathbb{R}} \circ T_a \circ (1_{\mathbb{R}})^{-1} = T_a$ (recall that R_0 and $1_{\mathbb{R}}$ are the only elements in $\langle R_0 \rangle$). Since $\langle R_0 \rangle$ is a group of order two, we may equate it with the cyclic group C_2 (by mapping $1_{\mathbb{R}}$ to 1 and R_0 to -1), which gives us $\mathrm{Iso}(\mathbb{R}) \cong C_2 \ltimes_\phi T(\mathbb{R})$ where $\phi_1(T_a) = T_a$ and $\phi_{-1}(T_a) = (T_a)^{-1}$. We can now see that ϕ is identical to the action in the definition of the dihedralization. Hence $\mathrm{Iso}(\mathbb{R}) \cong D(T(\mathbb{R}))$. □

Note that $\mathrm{Iso}(\mathbb{R})$ is also isomorphic to the dihedralization $D(\mathbb{R})$ since $T(\mathbb{R}) \cong \mathbb{R}$ (the function which sends T_a to a is an isomorphism).

Like $\mathrm{Iso}(\mathbb{R})$, $\mathrm{Aut}(\mathbb{D})$ is isomorphic to the dihedralization of its subgroup of translations (which we denote by $\tau(\mathbb{D})$). Again, we first show that $\tau(\mathbb{D})$ is indeed a subgroup, and in addition, that it is normal in $\mathrm{Aut}(\mathbb{D})$. We then show that $\mathrm{Aut}(\mathbb{D})$ is an internal semidirect product. Finally, we show that the externalization of this semidirect product is isomorphic to the dihedralization of $\tau(\mathbb{D})$.

Lemma 3.18. *The set $\tau(\mathbb{D})$ of all translations of the digital line is a subgroup of* $\mathrm{Aut}(\mathbb{D})$.

Proof. Let τ_a and τ_b be translations of the digital line. Since a and b are even, $a+b$ is even, and so $\tau_a \circ \tau_b = \tau_{a+b} \in \tau(\mathbb{D})$. Furthermore, $(\tau_a)^{-1} = \tau_{-a} \in \tau(\mathbb{D})$. Hence $\tau(\mathbb{D})$ is a subgroup of $\mathrm{Aut}(\mathbb{D})$. □

The subgroup $\tau(\mathbb{D})$ is also abelian since $\tau_a \circ \tau_b = \tau_{a+b} = \tau_b \circ \tau_a$.

Lemma 3.19. *The subgroup of translations $\tau(\mathbb{D})$ is normal in* $\mathrm{Aut}(\mathbb{D})$.

Proof. From the proof of Proposition 3.12, it is clear that $\tau(\mathbb{D})$ has index two in $\mathrm{Aut}(\mathbb{D})$. Since subgroups of index two are normal, $\tau(\mathbb{D})$ is normal in $\mathrm{Aut}(\mathbb{D})$. □

Lemma 3.20. *The automorphism group of the digital line admits an internal semidirect product decomposition* $\mathrm{Aut}(\mathbb{D}) = \langle \rho_0 \rangle \ltimes \tau(\mathbb{D})$.

Proof. Let $f \in \mathrm{Aut}(\mathbb{D})$. Then $f = \tau_n$ for some even n or $f = \rho_n$ for some (even or odd) n. If $f = \tau_n$, then $f = 1_{\mathbb{D}} \circ \tau_n$. If $f = \rho_n$, then $f = \rho_0 \circ \tau_{-2n}$ by Proposition 3.12. Hence $\mathrm{Aut}(\mathbb{D}) = \langle \rho_0 \rangle \tau(\mathbb{D})$. Since $\tau(\mathbb{D})$ is normal in $\mathrm{Aut}(\mathbb{D})$ and $\langle \rho_0 \rangle \cap \tau_{\mathbb{D}} = 1_{\mathbb{D}}$, we conclude that $\mathrm{Aut}(\mathbb{D}) = \langle \rho_0 \rangle \ltimes \tau(\mathbb{D})$. □

Theorem 3.21. *The automorphism group of the digital line is isomorphic to the dihedralization $D(\tau(\mathbb{D}))$.*

Proof. We have already proved that $\text{Aut}(\mathbb{D}) \cong \langle \rho_0 \rangle \ltimes \tau(\mathbb{D})$. We can write this internal semidirect product externally as $\langle \rho_0 \rangle \ltimes_\phi \tau(\mathbb{D})$ where $\phi_{\rho_0} : \tau(\mathbb{D}) \to \tau(\mathbb{D})$ is defined by $\phi_{\rho_0}(\tau_a) = \rho_0 \circ \tau_a \circ (\rho_0)^{-1} = (\tau_a)^{-1}$ and $\phi_{1_\mathbb{D}} : \tau(\mathbb{D}) \to \tau(\mathbb{D})$ is defined by $\phi_{1_\mathbb{D}}(\tau_a) = 1_\mathbb{D} \circ \tau_a \circ 1_\mathbb{D} = \tau_a$. Since $\langle \rho_0 \rangle$ is a group of order two, we may equate it with C_2 by mapping $1_\mathbb{D}$ to 1 and mapping ρ_0 to -1. Hence $\text{Aut}(\mathbb{D}) \cong C_2 \ltimes_\phi \tau(\mathbb{D})$ where $\phi_1(\tau_a) = \tau_a$ and $\phi_{-1}(\tau_a) = (\tau_a)^{-1}$. We now see that ϕ is identical to the action used in the definition of the dihedralization. Hence $\text{Aut}(\mathbb{D}) \cong D(\tau(\mathbb{D}))$. □

Note that $\text{Aut}(\mathbb{D})$ is also isomorphic to the dihedralization $D(\mathbb{Z})$ since $\tau(\mathbb{D}) \cong \mathbb{Z}$ (the function which sends τ_n to $\frac{n}{2}$ is an isomorphism).

4 Fundamental Groups

The next analogy we examine between the digital line and circles and their Euclidean counterparts involves a group called the *fundamental group*, which is one of the most basic invariants in algebraic topology. Intuitively, the fundamental group of a topological space X measures the extent to which special types of paths in X called *loops* can be deformed into one another. For example, the fundamental group of the torus is not isomorphic to the fundamental group of S^2: all loops in S^2 can be deformed into one another, but loops in the torus which wrap around the hole in the center cannot be deformed into loops which do not.

We begin by constructing the fundamental group of a *pointed topological space*. We then compute the fundamental groups of the real line and S^1 (which are standard results in algebraic topology). Finally, we show that the fundamental group of the real line is isomorphic to the fundamental group of the digital line, and that the fundamental group of S^1 is isomorphic to the fundamental group of each digital circle. This last proof will use continuous functions called *covering maps*, which are analogously constructed in the real and digital cases.

4.1 The Construction of the Fundamental Group

The fundamental group is defined for a slightly modified version of a topological space called a *pointed topological space*.

Definition 4.1. A *pointed topological space* (X, x_0) is a topological space X together with a *basepoint* $x_0 \in X$.

We can form a pointed space (X, x_0) from a space X by simply choosing a basepoint $x_0 \in X$. The definition of continuity for functions between pointed spaces is the same as that for functions between traditional topological spaces. However, we always require that functions between pointed spaces fix basepoints. That is, continuous functions $f : (X, x_0) \to (Y, y_0)$ have the additional requirement that $f(x_0) = y_0$.

We now introduce *loops*, which are the building blocks of the fundamental group.

Definition 4.2. A *loop* in a pointed space (X, x_0) is a path $\alpha : [0, 1] \to X$ with $\alpha(0) = \alpha(1) = x_0$. We say that α is *based* at x_0.

We are interested in the extent to which loops in a pointed space can be deformed into one another. The following definition is a precise statement of what we mean by "deform."

Definition 4.3. Let $f : X \to Y$ and $g : X \to Y$ be continuous functions. A *homotopy* deforming f into g is a continuous function $H : X \times [0, 1] \to Y$ such that $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$. If such a function exists, we say that f is *homotopic* to g and write $f \simeq g$.

We can think of a homotopy as a deformation of f into g throughout the time interval $[0, 1]$. Note that the homotopies in Definition 4.3 are deformations between arbitrary continuous functions $f : X \to Y$ and $g : X \to Y$. Since the fundamental group arises from deformations between loops, which are a type of path, we now introduce a modified version of a homotopy called a *path homotopy*.

Definition 4.4. Let $\alpha : [0, 1] \to X$ and $\beta : [0, 1] \to X$ be paths such that $\alpha(0) = \beta(0) = x_0$ and $\alpha(1) = \beta(1) = x_1$. A *path homotopy* deforming α into β is a homotopy $H : [0, 1] \times [0, 1] \to X$ between α and β with the additional requirement that $H(0, t) = x_0$ and $H(1, t) = x_1$ for all t. If such a function exists, we say that α is *path homotopic* to β and write $\alpha \simeq_p \beta$.

The additional requirement ensures that the restriction of a path homotopy $H : [0, 1]\times[0, 1] \to X$ to the subset $[0, 1] \times \{t\}$ is a path in X from x_0 to x_1 for any t.

It turns out that both $\simeq$ and $\simeq_p$ are equivalence relations. The proof is not difficult, and may be found in [4, p. 324]. This allows us to partition the loops in a pointed space (X, x_0) into equivalence classes of path homotopic loops. These equivalence classes are the elements of the fundamental group of (X, x_0). We also have a way to multiply certain paths together which we use to define an operation on equivalence classes of loops.

Definition 4.5. Let $\alpha : [0, 1] \to X$ and $\beta : [0, 1] \to X$ be paths with $\alpha(1) = \beta(0)$. The *product* of α and β is the path $\alpha * \beta : [0, 1] \to X$ defined by

$$(\alpha * \beta)(t) = \begin{cases} \alpha(2t) & \text{for } t \in [0, \frac{1}{2}] \\ \beta(2t - 1) & \text{for } t \in [\frac{1}{2}, 1]. \end{cases}$$

Note that the product $\alpha * \beta$ is continuous by the pasting lemma. If α is a path from x_0 to x_1, and β is a path from x_1 to x_2, then $\alpha * \beta$ is a path from x_0 to x_2. So if α and β are loops based at x_0, then $\alpha * \beta$ is also a loop based at x_0.

We now formally define the fundamental group of a pointed space (X, x_0).

Definition 4.6. Let (X, x_0) be a pointed space. The *fundamental group* of (X, x_0), written $\pi_1(X, x_0)$, is the set of equivalence classes of loops based at x_0 under the $\simeq_p$ relation with multiplication defined by $[\alpha][\beta] = [\alpha * \beta]$.

It is not immediately clear that the operation is well-defined since there is no unique representative for a given path homotopy class. The proof that the operation is well-defined is not difficult and may be found in [4, p. 326]. The identity element of the fundamental group is the homotopy class $[e]$ where $e : [0, 1] \to X$ is the constant loop defined by $e(t) = x_0$. The inverse of $[\alpha] \in \pi_1(X, x_0)$ is the class $[\bar{\alpha}]$ where $\bar{\alpha} : [0, 1] \to X$ is defined by $\bar{\alpha}(t) = \alpha(1 - t)$. Associativity holds as well, but is more difficult to prove. A proof of all these facts may be found in [4, p. 328].

Since the fundamental group is defined for a pointed space, $\pi_1(X, x_0)$ is generally not isomorphic to $\pi_1(X, x_1)$ when $x_0 \neq x_1$. However, these groups are isomorphic when X is path connected.

Proposition 4.7. *Let X be a space and let $x_0, x_1 \in X$. If X is path connected, then $\pi_1(X, x_0)$ is isomorphic to $\pi_1(X, x_1)$.*

A proof of Proposition 4.7 may be found in [4, p. 332]. So when computing the fundamental group of a path connected space X, we are free to choose any element of X as the basepoint. We may also write $\pi_1(X)$ instead of $\pi_1(X, x_0)$ and refer to *the* fundamental group of X. If X is a path connected space and $\pi_1(X)$ is trivial, we say that X is *simply connected*.

4.2 The Fundamental Groups of the Real Line and S^1

We now compute the fundamental groups of the real line and S^1. These are standard results in algebraic topology. The techniques we use to compute $\pi_1(S^1)$ will also be used to compute $\pi_1(C_n)$ for each even $n \geq 4$.

Theorem 4.8. *The fundamental group of the real line is trivial.*

Proof. Choose 0 as the basepoint in the real line. Let $[\alpha] \in \pi_1(\mathbb{R})$. The function $H : [0,1] \times [0,1] \to \mathbb{R}$ defined by

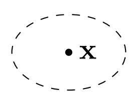
$$H(s,t) = (1-t)\alpha(s)$$

is clearly a path homotopy from α to the constant loop $e : [0,1] \to \mathbb{R}$ defined by $e(t) = 0$ (the constant loop at the basepoint). It follows that $[\alpha] = [e]$, and so $[e]$ is the only element in $\pi_1(\mathbb{R})$. □

It is usually quite difficult to compute the fundamental group of a space if that group is nontrivial. As we will see, S^1 and the digital circles have nontrivial fundamental groups. In order to compute them, we need to develop more complex tools. These tools are known as *covering spaces* and *lifting correspondences.*

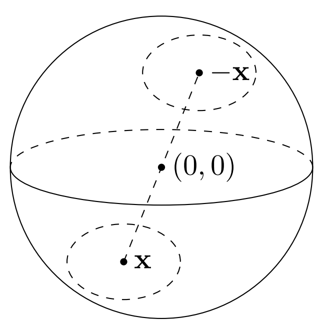
Definition 4.9. Let (X, x_0) and (Y, y_0) be pointed spaces. Let $p : (Y, y_0) \to (X, x_0)$ be a continuous surjection (with $p(y_0) = x_0$). If there exists an open cover $\{U_\alpha\}$ of X such that the inverse image $p^{-1}(U_\alpha)$ is a disjoint union of open sets for each α, and the restriction of p to each disjoint open set is a homeomorphism onto U_α, we say that p is a *covering map* from (Y, y_0) to (X, x_0). We call (X, x_0) the *base space* and (Y, y_0) the *covering space.*

Example 4.10. Recall that the *real projective plane* $\mathbb{R}\mathrm{P}^2$ can be defined as the quotient space of S^2 where $\mathbf{x} \sim -\mathbf{x}$ for all $\mathbf{x} \in S^2$. Although the projective plane cannot be embedded in $\mathbb{R}^3$ (and therefore cannot be visualized), it is a 2-manifold, which means that each $\mathbf{x} \in \mathbb{R}\mathrm{P}^2$ is contained in an open set which is homeomorphic to an open disk in $\mathbb{R}^2$. If $U_\mathbf{x}$ denotes such a set, then the collection $\{U_\mathbf{x} \mid \mathbf{x} \in \mathbb{R}\mathrm{P}^2\}$ is an open cover of $\mathbb{R}\mathrm{P}^2$.



An open set $U_\mathbf{x} \subset \mathbb{R}\mathrm{P}^2$ which is homeomorphic to an open disk in $\mathbb{R}^2$

Let $q : S^2 \to \mathbb{R}\mathrm{P}^2$ be the quotient map. Choose any $\mathbf{x}_0 \in S^2$ as a basepoint, and choose $q(\mathbf{x}_0)$ as the basepoint in $\mathbb{R}\mathrm{P}^2$. For each $\mathbf{x} \in \mathbb{R}\mathrm{P}^2$, the inverse image $q^{-1}(U_\mathbf{x})$ is a union of two open sets in S^2 lying opposite each other on the sphere:



$q^{-1}(U_\mathbf{x})$

We can also require that each $U_{\mathbf{x}}$ be small enough so that $q^{-1}(U_{\mathbf{x}})$ is a disjoint union of these open patches, as in the preceding figure. Since the restriction of q to each of these disjoint open sets is a homeomorphism onto $U_{\mathbf{x}}$, the quotient map q is a covering map, and so S^2 is a covering space of $\mathbb{R}\mathrm{P}^2$.

The following proposition lists one of the most important properties of covering spaces:

Proposition 4.11. *Let (X, x_0) and (Y, y_0) be pointed spaces, and let $p : (Y, y_0) \to (X, x_0)$ be a covering map. For each loop α in (X, x_0), there exists a unique path $\tilde{\alpha}$ in Y beginning at y_0 such that $p \circ \tilde{\alpha} = \alpha$.*

We say that the path $\tilde{\alpha}$ is a *lift* of α. Note that although α is a loop, $\tilde{\alpha}$ is not necessarily a loop. However, it is clear that $\tilde{\alpha}$ begins and ends in the set $p^{-1}(x_0)$. We call $p^{-1}(x_0)$ the *fiber* of p. Given a covering map, we now define a function from the fundamental group of the base space to the fiber.

Definition 4.12. Let (X, x_0) and (Y, y_0) be pointed spaces, and let $p : (Y, y_0) \to (X, x_0)$ be a covering map. The *lifting correspondence* associated with p is the function $\phi : \pi_1(X, x_0) \to p^{-1}(x_0)$ defined by $\phi([\alpha]) = \tilde{\alpha}(1)$ (where $\tilde{\alpha}$ is the unique lift from Proposition 4.11).

It is not immediately obvious that ϕ is well-defined, but this follows from the fact that we can lift not only loops, but also homotopies. A proof that ϕ is well-defined may be found in [4, p. 345]. If the covering space is simply connected, then the lifting correspondence is bijective. The existence of a bijective lifting correspondence will be our main tool for computing the fundamental groups of S^1 and the digital circles.

Proposition 4.13. *Let (X, x_0) and (Y, y_0) be pointed spaces, and let $p : (Y, y_0) \to (X, x_0)$ be a covering map. If Y is simply connected, then the lifting correspondence $\phi : \pi_1(X, x_0) \to p^{-1}(x_0)$ is bijective.*

A proof of Proposition 4.13 may be found in [4, p. 345].

Example 4.14. In Example 4.10, we showed that the quotient map $q : S^2 \to \mathbb{R}\mathrm{P}^2$ is a covering map. Let $\mathbf{x}_0 \in \mathbb{R}\mathrm{P}^2$ be a basepoint (we may choose any basepoint since the projective plane is path connected). The fiber $q^{-1}(\mathbf{x}_0)$ clearly has cardinality 2. Since S^2 is simply connected (a proof may be found in [4, p. 369]), the lifting correspondence $\phi : \pi_1(\mathbb{R}\mathrm{P}^2) \to q^{-1}(\mathbf{x}_0)$ associated with q is bijective, and so $\pi_1(\mathbb{R}\mathrm{P}^2)$ is a group of order two. Since all groups of order two are isomorphic to $\mathbb{Z}_2$, we conclude that $\pi_1(\mathbb{R}\mathrm{P}^2) \cong \mathbb{Z}_2$.

In order to compute $\pi_1(S^1)$, we first prove that the real line is a covering space of S^1.

Lemma 4.15. *The function $p : (\mathbb{R}, 0) \to (S^1, (1, 0))$ defined by $p(x) = (\cos 2\pi x, \sin 2\pi x)$ is a covering map.*

Proof. Clearly p is a continuous surjection. Consider the open cover of S^1 consisting of the following four open sets:

1. $\{(x, y) \in S^1 \mid x > 0\}$
2. $\{(x, y) \in S^1 \mid x < 0\}$
3. $\{(x, y) \in S^1 \mid y > 0\}$

4. $\{(x,y) \in S^1 \mid y < 0\}$.

We calculate the inverse image of each set in the cover as follows:

1. $p^{-1}(\{(x,y) \in S^1 \mid x > 0\}) = \{x \in \mathbb{R} \mid \cos 2\pi x > 0\} = \{(n - \frac{1}{4}, n + \frac{1}{4}) \mid n \in \mathbb{Z}\}$
2. $p^{-1}(\{(x,y) \in S^1 \mid x < 0\}) = \{x \in \mathbb{R} \mid \cos 2\pi x < 0\} = \{(n + \frac{1}{4}, n + \frac{3}{4}) \mid n \in \mathbb{Z}\}$
3. $p^{-1}(\{(x,y) \in S^1 \mid y > 0\}) = \{y \in \mathbb{R} \mid \sin 2\pi y > 0\} = \{(n, n + \frac{1}{2}) \mid n \in \mathbb{Z}\}$
4. $p^{-1}(\{(x,y) \in S^1 \mid y < 0\}) = \{y \in \mathbb{R} \mid \sin 2\pi y < 0\} = \{(n - \frac{1}{2}, n) \mid n \in \mathbb{Z}\}$.

Each of these inverse images is a disjoint union of open intervals, and the restriction of p to each interval is a bijection onto the original set in S^1. To see that this restriction is a homeomorphism, note that the restriction of p to the closure of one the intervals in the inverse image is a continuous bijection between compact Hausdorff spaces, and is therefore a homeomorphism. The further restriction of p to the interior of the closed interval (the original open interval we are interested in) is still a homeomorphism. □

For the next theorem, note that $p^{-1}((1,0)) = \mathbb{Z}$.

Theorem 4.16. *Let $p : (\mathbb{R}, 0) \to (S^1, (1,0))$ be the covering map from Lemma 4.15. The lifting correspondence $\phi : \pi_1(S^1, (1,0)) \to \mathbb{Z}$ associated with p is an isomorphism.*

Proof. Let $[\alpha], [\beta] \in \pi_1(S^1)$, and let $\tilde{\alpha}$ and $\tilde{\beta}$ be the unique lifts beginning at 0 of α and β respectively. Let $n = \tilde{\alpha}(1)$. Define a function $\gamma : [0,1] \to \mathbb{R}$ by $\gamma(t) = n + \tilde{\beta}(t)$ (this function is clearly continuous). Since $p(n) = \alpha(1)$, we know that $(\cos 2\pi n, \sin 2\pi n) = (1,0)$, and so $\cos 2\pi n = 1$ and $\sin 2\pi n = 0$. We now have

$$\begin{aligned}
p(\gamma(t)) &= p(n + \tilde{\beta}(t)) \\
&= (\cos 2\pi(n + \tilde{\beta}(t)), \sin 2\pi(n + \tilde{\beta}(t))) \\
&= (\cos 2\pi n \cos 2\pi\tilde{\beta}(t) - \sin 2\pi n \sin 2\pi\tilde{\beta}(t), \sin 2\pi n \cos 2\pi\tilde{\beta}(t) + \cos 2\pi n \sin 2\pi\tilde{\beta}(t)) \\
&= (\cos 2\pi\tilde{\beta}(t), \sin 2\pi\tilde{\beta}(t)) \\
&= p(\tilde{\beta}(t)) \\
&= \beta(t)
\end{aligned}$$

so γ is the lift of β beginning at $\gamma(0) = n + \tilde{\beta}(0) = n$. Since $\tilde{\alpha}(1) = \gamma(0)$, the product $\tilde{\alpha} * \gamma$ is defined. Additionally, we have

$$\begin{aligned}
p((\tilde{\alpha} * \gamma)(t)) &= (\cos 2\pi(\tilde{\alpha} * \gamma)(t), \sin 2\pi(\tilde{\alpha} * \gamma)(t)) \\
&= \begin{cases} (\cos 2\pi\tilde{\alpha}(2t), \sin 2\pi\tilde{\alpha}(2t)) & \text{if } t \in [0, \frac{1}{2}] \\ (\cos 2\pi\gamma(2t-1), \sin 2\pi\gamma(2t-1)) & \text{if } t \in [\frac{1}{2}, 1] \end{cases} \\
&= \begin{cases} p(\tilde{\alpha}(2t)) & \text{if } t \in [0, \frac{1}{2}] \\ p(\gamma(2t-1)) & \text{if } t \in [\frac{1}{2}, 1] \end{cases} \\
&= \begin{cases} \alpha(2t) & \text{if } t \in [0, \frac{1}{2}] \\ \beta(2t-1) & \text{if } t \in [\frac{1}{2}, 1] \end{cases} \\
&= (\alpha * \beta)(t)
\end{aligned}$$

so $\tilde{\alpha} * \gamma$ is the unique lift of $\alpha * \beta$ beginning at 0. Finally,

$$\begin{aligned}\phi([\alpha * \beta]) &= (\tilde{\alpha} * \gamma)(1) \\ &= \gamma(1) \\ &= n + \tilde{\beta}(1) \\ &= \phi([\alpha]) + \phi([\beta])\end{aligned}$$

so ϕ is a homomorphism. Since the real line is simply connected, we conclude from Proposition 4.13 that ϕ is an isomorphism. □

In other words, the fundamental group of S^1 is isomorphic to the integers under addition.

4.3 The Fundamental Groups of the Digital Line and Circles

We now show that the digital line and real line have isomorphic fundamental groups, as do each digital circle and S^1. That is, we show that $\pi_1(\mathbb{D})$ is trivial and $\pi_1(C_n) \cong \mathbb{Z}$ for each even $n \geq 4$.

As in our computation of $\pi_1(\mathbb{R})$, we compute $\pi_1(\mathbb{D})$ by showing that any loop is homotopic to the constant loop at the basepoint. First, we must prove that the image of a loop in the digital line is a digital interval, that is, a finite set of consecutive integers.

Lemma 4.17. *The image of a loop in the digital line is a finite set of consecutive integers.*

Proof. Let $\alpha : [0, 1] \to \mathbb{D}$ be a loop. Since $[0, 1]$ is compact and connected, $\text{im}(\alpha)$ is compact and connected. It follows from Propositions 2.11 and 2.12 that $\text{im}(\alpha)$ is a finite set of consecutive integers. □

The fact that the image of a loop $\alpha : [0, 1] \to \mathbb{D}$ is a digital interval will allow us to construct a series of homotopies from α to the constant loop at the basepoint. To do this, we will construct a path β whose image is one size smaller than $\text{im}(\alpha)$ and show that $\alpha \simeq_p \beta$. We then iterate this argument to show that α is homotopic to a loop whose image is a one-point set, that is, the constant loop at the basepoint.

Theorem 4.18. *The fundamental group of the digital line is trivial.*

Proof. First, choose 0 as the basepoint. Let $\alpha : [0, 1] \to \mathbb{D}$ be a non-constant loop based at 0. We know from Lemma 4.17 that $\text{im}(\alpha)$ is a finite set of consecutive integers. Since α is non-constant, $|\text{im}(\alpha)| \geq 2$ and has a maximum and a minimum, one of which is not equal to 0. Let $|\text{im}(\alpha)| = k$. We will show that α is path homotopic to a loop $\beta : [0, 1] \to \mathbb{D}$ based at 0 with $|\text{im}(\beta)| = k - 1$. The proof has four cases:

1. $\max(\text{im}(\alpha)) \neq 0$ and is odd;
2. $\max(\text{im}(\alpha)) \neq 0$ and is even;
3. $\min(\text{im}(\alpha)) \neq 0$ and is odd;
4. $\min(\text{im}(\alpha)) \neq 0$ and is even.

We give proofs for the first and second cases. The proof of the third case is similar to the first, and the proof of the fourth case is similar to the second.

Suppose for the first case that $\max(\text{im}(\alpha)) \neq 0$ and is odd. Let $m = \max(\text{im}(\alpha))$. Let $\beta : [0,1] \to \mathbb{D}$ be the function defined by

$$\beta(t) = \begin{cases} \alpha(t) & \text{if } \alpha(t) \neq m \\ m-1 & \text{if } \alpha(t) = m. \end{cases}$$

Note that $\text{im}(\beta) = \text{im}(\alpha) \setminus \{m\}$, and so $|\text{im}(\beta)| = k-1$. For $n \leq m-2$, it is a consequence of the digital line's topology that $m-1 \notin B(n)$. So $\beta^{-1}(B(n)) = \alpha^{-1}(B(n))$ is open since α is continuous. For $B(m-1) = \{m-2, m-1, m\}$, we have

$$\begin{aligned} \beta^{-1}(B(m-1)) &= \beta^{-1}(\{m-2, m-1, m\}) \\ &= \beta^{-1}(\{m-2, m-1\}) \cup \beta^{-1}(\{m\}) \\ &= \beta^{-1}(\{m-2, m-1\}) \cup \varnothing \\ &= \alpha^{-1}(\{m-2, m-1, m\}) \\ &= \alpha^{-1}(B(m-1)), \end{aligned}$$

which is also open since α is continuous. Since inverse images of basis elements are open, β is continuous. Since $\beta(0) = \alpha(0) = 0$ and $\beta(1) = \alpha(1) = 0$, β is a loop based at 0. Now let $H : [0,1] \times [0,1] \to \mathbb{D}$ be the function defined by

$$H(s,t) = \begin{cases} \alpha(s) & \text{if } t \in [0,1) \\ \beta(s) & \text{if } t = 1. \end{cases}$$

We know from the above calculations that $\beta^{-1}(B(n)) = \alpha^{-1}(B(n))$ when $n \leq m-1$. So for $n \leq m-1$, we have

$$\begin{aligned} H^{-1}(B(n)) &= \alpha^{-1}(B(n)) \times [0,1) \cup \beta^{-1}(B(n)) \times \{1\} \\ &= \alpha^{-1}(B(n)) \times [0,1) \cup \alpha^{-1}(B(n)) \times \{1\} \\ &= \alpha^{-1}(B(n)) \times [0,1], \end{aligned}$$

which is open since α is continuous. The last basis element we need to check is $B(m) = \{m\}$. Since $\beta^{-1}(\{m\}) = \varnothing$, we have

$$\begin{aligned} H^{-1}(B(m)) &= \alpha^{-1}(B(m)) \times [0,1) \cup \beta^{-1}(B(m)) \times \{1\} \\ &= \alpha^{-1}(\{m\}) \times [0,1) \cup \varnothing \\ &= \alpha^{-1}(\{m\}) \times [0,1), \end{aligned}$$

which is also an open set. We now know that H is continuous. Since $H(s,0) = \alpha(s)$, $H(s,1) = \beta(s)$, and $H(0,t) = H(1,t) = 0$, we conclude that H is a path homotopy deforming α into β. This concludes case one.

Suppose for the second case that $\max(\text{im}(\alpha)) \neq 0$ and is even. Let $m = \max(\text{im}(\alpha))$. Let $\beta : [0,1] \to \mathbb{D}$ be defined as in the first case. Since our argument for the continuity of β depended on the fact that m was odd, we must now show that β is continuous when m is even. For $n \leq m-3$, $m-1 \notin B(n)$, and so $\beta^{-1}(B(n)) = \alpha^{-1}(B(n))$ is open. Since $m-2$ is even, $B(m-2) = \{m-3, m-2, m-1\}$. Note that $\beta^{-1}(\{m-1\}) = \alpha^{-1}(\{m-1, m\})$ and $\alpha^{-1}(m+1) = \varnothing$ since

$m = \max(\mathrm{im}(\alpha))$. So we have

$$\begin{aligned}\beta^{-1}(B(m-2)) &= \beta^{-1}(\{m-3, m-2, m-1\}) \\ &= \alpha^{-1}(\{m-3, m-2, m-1, m\} \\ &= \alpha^{-1}(\{m-3, m-2, m-1, m, m+1\} \\ &= \alpha^{-1}(\{m-3, m-2, m-1\} \cup \alpha^{-1}(\{m-1, m, m+1\} \\ &= \alpha^{-1}(B(m-2)) \cup \alpha^{-1}(B(m)),\end{aligned}$$

which is a union of two open sets, and therefore open. Similarly we have that

$$\begin{aligned}\beta^{-1}(B(m-1)) &= \beta^{-1}(\{m-1\}) \\ &= \alpha^{-1}(\{m-1, m\}) \\ &= \alpha^{-1}(\{m-1, m, m+1\}) \\ &= \alpha^{-1}(B(m))\end{aligned}$$

is open. Finally, since $m \notin \mathrm{im}(\beta)$ and $m+1 \notin \mathrm{im}(\beta)$, we have

$$\begin{aligned}\beta^{-1}(B(m)) &= \beta^{-1}(\{m-1, m, m+1\}) \\ &= \beta^{-1}(\{m-1\}) \\ &= \alpha^{-1}(\{m-1, m\} \\ &= \alpha^{-1}(\{m-1, m, m+1\} \\ &= \alpha^{-1}(B(m)).\end{aligned}$$

Hence β is continuous. Now let $K : [0,1] \times [0,1] \to \mathbb{D}$ be the function defined by

$$K(s,t) = \begin{cases} \alpha(s) & \text{if } t = 0 \\ \beta(s) & \text{if } t \in (0,1]. \end{cases}$$

For $n \leq m-3$, we have

$$\begin{aligned}K^{-1}(B(n)) &= \alpha^{-1}(B(n)) \times \{0\} \cup \beta^{-1}(B(n)) \times (0,1] \\ &= \alpha^{-1}(B(n)) \times \{0\} \cup \alpha^{-1}(B(n)) \times (0,1] \\ &= \alpha^{-1}(B(n)) \times [0,1],\end{aligned}$$

which is open. Next, we have

$$\begin{aligned}K^{-1}(B(m-2)) &= K^{-1}(\{m-3, m-2, m-1\}) \\ &= \alpha^{-1}(\{m-3, m-2, m-1\}) \times \{0\} \cup \beta^{-1}(\{m-3, m-2, m-1\}) \times (0,1] \\ &= \alpha^{-1}(\{m-3, m-2, m-1\}) \times \{0\} \cup \alpha^{-1}(\{m-3, m-2, m-1, m\}) \times (0,1] \\ &= \alpha^{-1}(\{m-3, m-2, m-1\}) \times [0,1] \cup \alpha^{-1}(\{m-1, m\}) \times (0,1] \\ &= \alpha^{-1}(\{m-3, m-2, m-1\}) \times [0,1] \cup \alpha^{-1}(\{m-1, m, m+1\}) \times (0,1] \\ &= \alpha^{-1}(B(m-2)) \times [0,1] \cup \alpha^{-1}(B(m)) \times (0,1],\end{aligned}$$

which is open. Next, we have

$$\begin{aligned}
K^{-1}(B(m-1)) &= K^{-1}(\{m-1\}) \\
&= \alpha^{-1}(\{m-1\}) \times \{0\} \cup \beta^{-1}(\{m-1\}) \times (0,1] \\
&= \alpha^{-1}(\{m-1\}) \times \{0\} \cup \alpha^{-1}(\{m-1, m\}) \times (0,1] \\
&= \alpha^{-1}(\{m-1\}) \times [0,1] \cup \alpha^{-1}(\{m-1, m, m+1\}) \times (0,1] \\
&= \alpha^{-1}(B(m-1)) \times [0,1] \cup \alpha^{-1}(B(m)) \times (0,1],
\end{aligned}$$

which is also open. Finally,

$$\begin{aligned}
K^{-1}(B(m)) &= K^{-1}(\{m-1, m, m+1\}) \\
&= \alpha^{-1}(\{m-1, m, m+1\}) \times \{0\} \cup \beta^{-1}(\{m-1, m, m+1\}) \times (0,1] \\
&= \alpha^{-1}(\{m-1, m, m+1\}) \times \{0\} \cup \alpha^{-1}(\{m-1, m, m+1\}) \times (0,1] \\
&= \alpha^{-1}(\{m-1, m, m+1\}) \times [0,1] \\
&= \alpha^{-1}(B(m)) \times [0,1]
\end{aligned}$$

is open. Hence K is continuous and defines a homotopy deforming α into β. We now conclude case two. The third case is similar to the first, and the fourth case is similar to the second, both with a slight adjustment in the way we define β. Specifically, we send t to $m+1$ when $\alpha(t) = m$, where in this case $m = \min(\text{im}(\alpha))$.

We have now shown that any loop $\alpha : [0,1] \to \mathbb{D}$ based at 0 with $|\text{im}(\alpha)| = k$ is path homotopic to a loop $\beta : [0,1] \to \mathbb{D}$ based at 0 with $|\text{im}(\beta)| = k-1$. Since $\simeq_p$ is an equivalence relation, it follows immediately that α is path homotopic to the constant loop at 0 (we simply iterate the above arguments until $|\text{im}(\beta)| = 1$). Hence all loops based at 0 are path homotopic, and so $\pi_1(\mathbb{D})$ is trivial. □

To compute the fundamental groups of the digital circles, will use the same method as in our computation of $\pi_1(S^1)$. First, we show that the digital line covers each digital circle, just as the real line covers S^1.

Lemma 4.19. *The map $q : (\mathbb{D}, 0) \to (C_n, [2])$ defined by $q(x) = [(x \bmod n) + 2]$ is a covering map for each even $n \geq 4$.*

Proof. Consider the open cover consisting of the basis elements $B([x])$ for each $[x] \in C_n$. Note that

$$q^{-1}([x]) = \bigcup_{k \in \mathbb{Z}} \{nk + x - 2\}.$$

So if x is odd,

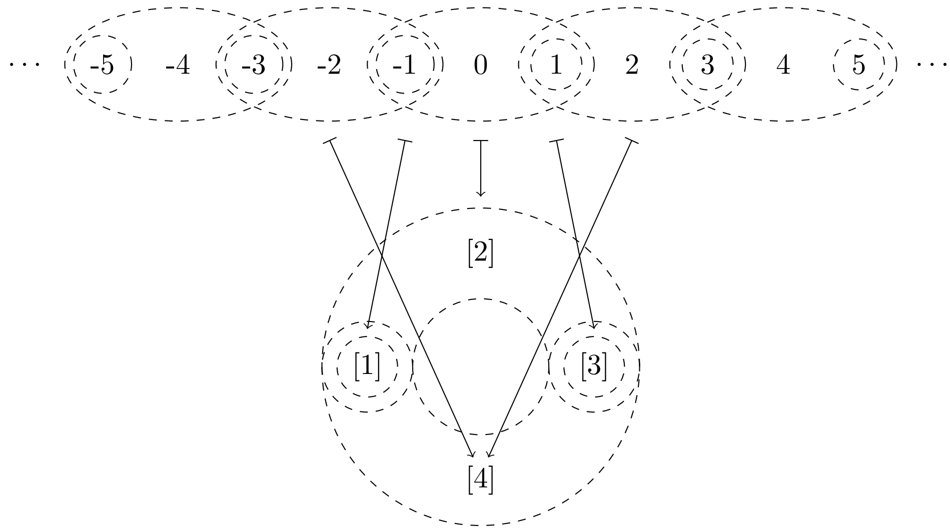
$$\begin{aligned}
q^{-1}(B([x])) &= q^{-1}(\{[x]\}) \\
&= \bigcup_{k \in \mathbb{Z}} \{nk + x - 2\}.
\end{aligned}$$

This set is a disjoint union of open sets since $nk + x + 2$ is odd for all values of k (this results from the fact that n is even and x is odd). If x is even,

$$\begin{aligned}
q^{-1}(B([x])) &= q^{-1}(\{[x-1], [x], [x+1]\}) \\
&= \bigcup_{k \in \mathbb{Z}} \{nk + x - 3, nk + x - 2, nk + x - 1\}
\end{aligned}$$

which is also a disjoint union of open sets since $nk + x + 2$ is even for all values of k (this results from the fact that both n and x are even). The restriction of q to each disjoint open set is also a homeomorphism since it is a continuous bijection and both the domain and codomain consist of a single open set. □

We can visualize the covering map from Lemma 4.19 as follows:



A portion of the digital line covering C_4

Note that this covering map is analogous the covering map $p : \mathbb{R} \to S^1$ of Lemma 4.15. While p is periodic, q is modular. Both maps may be thought of as wrapping the covering spaces around the base spaces. Like the real line, the digital line is simply connected, and so we have a bijective lifting correspondence from $\pi_1(C_n)$ (for each even $n \geq 4$) to the fiber $q^{-1}([2])$. We now show that this correspondence is a homomorphism, and therefore an isomorphism. Note that $q^{-1}([2]) = n\mathbb{Z}$.

Theorem 4.20. *Let $q : (\mathbb{D}, 0) \to (C_n, [2])$ be the covering map from Lemma 4.19. The lifting correspondence $\phi : \pi_1(C_n) \to n\mathbb{Z}$ associated with q is an isomorphism for each even $n \geq 4$.*

Proof. Note that the construction of our covering map uses [2] as the basepoint in each digital circle. Let $[\alpha], [\beta] \in \pi_1(C_n)$, and let $\tilde{\alpha}$ and $\tilde{\beta}$ be the liftings beginning at 0 of α and β respectively. Let $\gamma : [0, 1] \to \mathbb{D}$ be the function defined by $\gamma(t) = \tilde{\alpha}(1) + \tilde{\beta}(t)$. We know from the proof of Lemma 4.19 that $q^{-1}(B([2]))$ is a disjoint union of open sets. Furthermore, since $\tilde{\alpha}(1)$ is in the fiber of q, $\tilde{\alpha}(1)$ is contained in one of these open sets (call it U). Let r denote the restriction of q to U. We know that this function is a homeomorphism onto $B([2])$, and therefore a bijection. This means that $r^{-1}([2]) = \tilde{\alpha}(1)$. Since the restriction is continuous and $\{[2]\}$ is closed in each C_n, $r^{-1}(\{[2]\}) = \{\tilde{\alpha}(1)\}$ is closed in the digital line. It follows that $\tilde{\alpha}(1)$ is an even integer. Recall that for any even integer a, the translation function τ_a defined by $\tau_a(x) = x + a$ is continuous. We can rewrite γ as the composition of the translation function τ_m where $m = \tilde{\alpha}(1)$ and the loop $\tilde{\beta}$. That is, $\gamma = \tau_m \circ \tilde{\beta}$. Thus γ is continuous since it is the composition of two continuous functions. To proceed, note that $\tilde{\alpha}(1) \equiv 0 \pmod{n}$ since

$$\begin{aligned} p(\tilde{\alpha}(1)) = \alpha(1) &\implies [(\tilde{\alpha}(1) \bmod n) + 2] = [2] \\ &\implies \tilde{\alpha}(1) \bmod n = 0. \end{aligned}$$

This allows us to conclude that γ is a lifting of β since

$$\begin{aligned}
p(\gamma(t)) &= p(\tilde{\alpha}(1) + \tilde{\beta}(t)) \\
&= [((\tilde{\alpha}(1) + \tilde{\beta}(t)) \bmod n) + 2] \\
&= [(\tilde{\beta}(t) \bmod n) + 2] \\
&= p(\tilde{\beta}(t)) \\
&= \beta(t).
\end{aligned}$$

Furthermore, $\gamma(0) = \tilde{\alpha}(1) + \tilde{\beta}(0) = \tilde{\alpha}(1)$, so $\tilde{\alpha} * \gamma$ is defined. From the definition of the path product, we see that

$$\begin{aligned}
p((\tilde{\alpha} * \gamma)(t)) &= \begin{cases} p(\tilde{\alpha}(2t)) & \text{if } t \in [0, \frac{1}{2}] \\ p(\gamma(2t-1)) & \text{if } t \in [\frac{1}{2}, 1] \end{cases} \\
&= \begin{cases} \alpha(2t) & \text{if } t \in [0, \frac{1}{2}] \\ \beta(2t-1) & \text{if } t \in [\frac{1}{2}, 1] \end{cases} \\
&= (\alpha * \beta)(t)
\end{aligned}$$

so $\tilde{\alpha} * \gamma$ is the lifting of $\alpha * \beta$ beginning at 0. Since

$$\begin{aligned}
\phi\big([\alpha][\beta]\big) &= \phi\big([\alpha * \beta]\big) \\
&= (\tilde{\alpha} * \gamma)(1) \\
&= \gamma(1) \\
&= \tilde{\alpha}(1) + \tilde{\beta}(1) \\
&= \phi\big([\alpha]\big) + \phi\big([\beta]\big)
\end{aligned}$$

we conclude that the lifting correspondence is a homomorphism, and thus an isomorphism, for each even $n \geq 4$. □

In other words, $\pi_1(C_n) \cong n\mathbb{Z}$ for each even $n \geq 4$. Since $n\mathbb{Z} \cong \mathbb{Z}$, we conclude that $\pi_1(C_n) \cong \mathbb{Z}$ for each even $n \geq 4$. We have now shown that the digital line and circles and their Euclidean counterparts have isomorphic fundamental groups. That is, $\pi_1(\mathbb{D})$ and $\pi_1(\mathbb{R})$ are trivial, and $\pi_1(C_n) \cong \pi_1(S^1) \cong \mathbb{Z}$ for each even $n \geq 4$.

5 Conclusion

We have now established four analogies between the real and digital lines and circles:

1. The groups $\mathrm{Iso}(\mathbb{R})$ and $\mathrm{Aut}(\mathbb{D})$ are both isomorphic to the dihedralizations of their subgroup of translations;

2. Both the real and digital lines are simply connected;

3. There exist analogously constructed covering maps from the real line to S^1 and from the digital line to each digital circle;

4. Both $\pi_1(S^1)$ and $\pi_1(C_n)$ (for each even $n \geq 4$) are isomorphic to $\mathbb{Z}$.

Furthermore, our fundamental group computations were accomplished without using simplicial methods or posets (which are often used to study finite spaces). In other words, we can "do topology" on the digital lines and circles with the same techniques we use to "do topology" on the real line and S^1. However, two important questions remain:

1. Is the digital line contractible?

2. Are the higher homotopy groups of each digital circle trivial?

It is well-known that the real line is contractible and that the higher homotopy groups of S^1 are trivial. Affirmatives answers to the above two questions would help make our analogies between the real and digital lines and circles more complete.

References

[1] Colin Adams and Robert Franzosa, *Introduction to topology: Pure and applied*, Pearson Prentice Hall, 2008.

[2] Barbara A. Brown, *Generalized dihedral groups of small order*, Undergraduate thesis, Available in Simpson Library.

[3] David S. Dummit and Richard M. Foote, *Abstract algebra*, John Wiley and Sons, 2004.

[4] James R. Munkres, *Topology*, Prentice Hall, 2000.

[5] John S. Rose, *A course on group theory*, Dover Publications, 1994.

[6] R. Kopperman T.Y. Kong and P.R. Meyer, *A topological approach to digital topology*, American Mathematical Monthly **98** (1991), no. 10, 901–917.