APPENDIX C
CHAPTER-SPECIFIC DETAILS

C.1 CHAPTER 1: HAT COLORS AND HAMMING CODES

The construction of Hamming codes for $n = 2^m - 1$, $m \in \mathbb{N}$, with $m > 2$ is maybe easiest when considering linear codes. To do so, we consider codewords such as (010) as an element in $\mathbb{Z}_2^3$ and will take their sums, for example. Here, writing $\mathbb{Z}_2$ is short for $\mathbb{Z}/2\mathbb{Z}$, which works exactly as the cyclic group $\mathbb{Z}/3\mathbb{Z}$ that we discuss in chapter 4. It means that we compute everything modulo 2. For example,

$$(010) + (100) = (110) \quad \text{or} \quad (011) + (111) + (001) = (101).$$

We can also multiply the sequences by 0 or 1:

$$0 \cdot (011) = (000) \quad \text{and} \quad 1 \cdot (011) = (011).$$

This is not a very exciting operation, but it allows us to consider the space $\mathbb{Z}_2^3$ as a so-called "linear space."

More importantly, we call a code W (that is, the collection of codewords) a *linear code* if W is a linear subspace of $\mathbb{Z}_2^n$. Or, equivalently, if the all-zero sequence is in W and if for all $x, y \in W$, we have $x + y \in W$, then W is a linear code. Later in this appendix we will introduce Hamming codes as linear codes. But before doing so, we introduce a few concepts from linear algebra that

will make our life much easier when constructing the Hamming codes and analyzing their properties.

Linear Algebra on $\mathbb{Z}_2^n$

We first show how matrix multiplication can work in the type of linear space we consider here. It is almost the same as in the real-valued case introduced in appendix B.7, except for that here, everything is computed modulo 2. In short, for a $p \times q$ matrix A and a $q \times r$ matrix B, the product $A \cdot B$ is a $p \times r$ matrix, whose (i, j)th entry is defined as

$$(A \cdot B)_{ij} := \sum_{k=1}^{q} A_{ik} B_{kj} \bmod 2.$$

Let us consider an example. To multiply the 1×7 matrix (or vector)

$$x := \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}$$

with the 7×3 matrix

$$H := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \tag{C.1}$$

it is easiest to consider one column of H at a time. We have

$$x \cdot H = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}.$$

Here, we computed the first element as

$$1 \cdot 0 + 0 \cdot 0 + 1 \cdot 0 + 0 \cdot 1 + 0 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 2 = 0 \bmod 2,$$

the second element as

$$1 \cdot 0 + 0 \cdot 1 + 1 \cdot 1 + 0 \cdot 0 + 0 \cdot 0 + 1 \cdot 1 + 1 \cdot 1 = 3 = 1 \bmod 2,$$

and the third element as

$$1 \cdot 1 + 0 \cdot 0 + 1 \cdot 1 + 0 \cdot 0 + 0 \cdot 1 + 1 \cdot 0 + 1 \cdot 1 = 3 = 1 \bmod 2.$$

The matrix H thus defines a map from $\mathbb{Z}_2^7$ to $\mathbb{Z}_2^3$, via $x \mapsto x \cdot H$. It is usually referred to as a *linear map*, because

$$(x + y) \cdot H = x \cdot H + y \cdot H.$$

We need two more concepts: linear independence of vectors and the dimension of a linear space. We call a set of vectors $x_1, \ldots, x_p \in \mathbb{Z}_2^n$ *linearly independent* if for $\lambda_1, \ldots, \lambda_p \in \mathbb{Z}_2$, we have

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_p x_p = 0 \quad \Longrightarrow \quad \lambda_1 = \lambda_2 = \cdots = \lambda_p = 0,$$

where, again, all computations are performed modulo 2. In words, there is only one way of combining the vectors $x_1, \ldots, x_p$ to get 0: We have to multiply each vector by 0. As an example, the vectors

$$x_1 = (1001),\ x_2 = (0101),\ x_3 = (1011),\ x_4 = (0010)$$

are not linearly independent (we then call them *linearly dependent*), since

$$1 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 + 1 \cdot x_4 = 0.$$

Given a a subset W of $\mathbb{Z}_2^n$ (e.g., our code), W is a *linear subspace* if the all-zero sequence is in W and

$$x, y \in W \quad \Longrightarrow \quad x + y \in W.$$

We can now define the *dimension* $\dim W$ of W. It is the maximum number of linearly independent vectors.[1] For example, the dimension of $\mathbb{Z}_2^4$ itself equals $\dim \mathbb{Z}_2^4 = 4$. This is the case,

1. Strictly speaking, one has to argue that this is a well-defined number by showing that any maximal set of linearly independent vectors has the same size. We are not going to prove that here, but we hope that you believe that this is the case.

because the unit vectors

$$e_1 := (0001),\ e_2 := (0010),\ e_3 := (0100),\ e_4 := (1000)$$

are linearly independent, and adding any other vector element would make the set linearly dependent (you may want to prove why this is the case). These maximal sets of linearly independent vectors are very practical. For example, one can show that each vector in the linear space can be written as a unique linear combination of such unit vectors (can you prove that, too?). For example, in $\mathbb{Z}_2^4$, the vector (1101) has the decomposition

$$1 \cdot e_1 + 1 \cdot e_2 + 0 \cdot e_3 + 1 \cdot e_4.$$

Among other things, this allows us to count the number of elements in W. Each factor can be either 0 or 1 and thus

$$\#W := 2^{\dim W}, \tag{C.2}$$

for example, $\#\mathbb{Z}_2^4 = 16$, which we know already, of course.

Construction of General Hamming Codes

Equipped with these tools, we can now define the Hamming code for $n = 2^m - 1$ using the $n \times m$ matrix

$$H := \begin{pmatrix} 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 & 1 \\ 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 1 & 0 \\ 0 & \cdots & 0 & 1 & 1 & 1 \\ 0 & \cdots & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \cdots & 1 & 1 & 1 & 0 \\ 1 & \cdots & 1 & 1 & 1 & 1 \end{pmatrix}.$$

It is constructed by concatenating all $n = 2^m - 1$ binary words of length m as rows, while excluding the all-zero sequence. We

define the code W as

$$W := \{x \in \mathbb{Z}_2^n : x \cdot H = 0\}.$$

Here, 0 stands for the all-zero sequence $(0 \cdots 0)$ that consists of n 0s; we can use this notation, since it is always clear from the context whether we mean the number $0 \in \mathbb{R}$ or the vector $0 \in \mathbb{Z}_2^n$. Equation (C.1) shows H for $n = 7$ and $m = 3$. In fact, the corresponding W is the Hamming code that forms the basis for the one we have shown in the figure on page 16. There we did not display the original codewords, but added the constant word (0110100) to avoid codewords like (0000000). Codeword 2 in the figure, for example, equals (0111011). And, indeed, for $x_2 = (0111011) - (0110100) = (0001111)$, we have $x_2 \cdot H = (0, 0, 0)$. How do we know that W is indeed a Hamming code? To prove this, we will exploit the fact that X is a linear code: $(x + y) \cdot H = x \cdot H + y \cdot H = 0$. This should not come as a surprise, given that we have gone through all this work!

Let us first prove that W is 1-error correcting. To do so, we must show that the Hamming distance between two codewords is at least 3:

$$\min_{x,y \in W} d(x, y) \geq 3.$$

Because W is linear and $d(0, y + x) = d(x + x, y + x) = d(x, y)$, this is equivalent to

$$\min_{y \in W} d(0, y) \geq 3.$$

Furthermore, $\min_{y \in W} d(0, y)$ equals the minimal number of linearly dependent rows in H. (This is shown by the following argument. Assume that the minimum is achieved for $y' \in W$. Then, y' has $d(0, y')$ nonzero entries, and since $y'H = 0$, the corresponding rows in H are linearly dependent. Conversely, a minimal set of linearly dependent rows can be represented by the ones in a vector $y' \in \mathbb{Z}_2$ that satisfies $y'H = 0$.) But clearly, the minimal number of linearly dependent rows of H equals 3: any

2 rows are linearly independent (no 2 rows are identical), and the first 3 rows are linearly dependent.

Finally, to prove that the code is perfect, we can again use a counting argument. We will show that each sequence has at most distance 1 to any of the codewords (that is, the balls with radius 1 span the whole space):

$$\bigcup_{x\in W} B_1(x) = \mathbb{Z}_2^n.$$

Note that we have[2] $\dim W = n - m$. We then have, using equation (C.2),

$$\# \bigcup_{x\in W} B_1(x) = \#W \cdot (n+1) = 2^{\dim W} \cdot 2^m = 2^{n-m} \cdot 2^m = 2^n = \#\mathbb{Z}_2^n,$$

which proves that the spaces must be equal.

The Hamming code W contains 2^{n-m} codewords. If the colors of the hats are distributed randomly, the probability of losing is, as stated in section 1.5,

$$P(\text{group loses}) = \frac{2^{n-m}}{2^n} = \frac{1}{2^m} = \frac{1}{n+1}.$$

The more players, the merrier!

C.2 CHAPTER 4: ANIMAL STICKERS AND CYCLIC GROUPS

Variation: Colored Hats with Infinitely Many Players

It is a mathematical curiosity that the colored-hats game still works with a countably infinite number of players and a finite number m of colors. This is, of course, irrelevant for all practical purposes, but we encourage you to indulge yourself with

2. The easiest way to see this is to use the rank-nullity theorem. It states that $\dim \mathbb{Z}_2^n = \dim \operatorname{im} H^\top + \dim \ker H^\top$, where $H^\top$ is the transpose of H, $\ker H^\top = W$ the kernel, and $\operatorname{im} H^\top$ is the image of $H^\top$. The dimension of $\operatorname{im} H^\top$ equals m, since H contains the linearly independent unit vectors e_i, $i = 1, \ldots, m$, as rows; that is, $H^\top$ has full column rank, in this case, m.

this thought experiment. It allows us to take a glimpse at the astonishing nature of the concept of infinity and the difficulties it raises.

Let us number the players with an increasing sequence, starting from 1 (the player in the back). The distribution of colors can then be represented by a sequence $(C_1, C_2, \ldots) = (C_k)_{k\in\mathbb{N}}$ of numbers $C_k \in \{0, \ldots, p-1\}$. It is now possible to construct an equivalence relation on the set of all such sequences:

$$(C_k)_{k\in\mathbb{N}} \sim (D_k)_{k\in\mathbb{N}} :\Leftrightarrow \text{only finitely many numbers are different.}$$

You can certainly convince yourself easily that this indeed defines an equivalence relation (which, in fact, is quite famous). Earlier we wrote down equivalence classes using so-called representatives. In $\mathbb{Z}/3\mathbb{Z}$, we wrote

$$\{\ldots, -1, 2, 5, 8, 11, 14, \ldots\} = [2],$$

for example; that is, we chose 2 as a representative of the corresponding equivalence class. Now, we will do the same for the infinite sequences and represent the (infinitely many) equivalence classes by

$$(C_k^1)_{k\in\mathbb{N}}, (C_k^2)_{k\in\mathbb{N}}, (C_k^3)_{k\in\mathbb{N}}, \ldots.$$

This seems like a natural and harmless step, but it is far from being that. To guarantee that such a choice is possible requires the so-called "axiom of choice," which has been the topic of many debates and mathematical paradoxes. When encountered for the first time, it can surprise you by showing how much mathematics is beyond the classes of "correct" and "incorrect."

For now, let us assume that we do have access to the representatives $(C_k^1)_{k\in\mathbb{N}}, (C_k^2)_{k\in\mathbb{N}}, (C_k^3)_{k\in\mathbb{N}}, \ldots$. Suppose that the current game is described by the true sequence $(C_k)_{k\in\mathbb{N}}$. The key observation for solving the (imaginary) game is that each player does not know the true sequence $(C_k)_{k\in\mathbb{N}}$ but recognizes its equivalence class,

$$[(C_k)_{k\in\mathbb{N}}] = [(C_k^{480213})_{k\in\mathbb{N}}],$$

say. Since $(C_k)_{k\in\mathbb{N}}$ and $(C_k^{480213})_{k\in\mathbb{N}}$ differ by only finitely many numbers, there is a number M, say, beyond which they are identical. The key idea is that any player a for whom the remaining sequence of hat colors $(C_k)_{k\geq a+1}$ they see matches the sequence of the representative $(C_k^{480213})_{k\geq a+1}$ of the equivalence class, guesses his own color to be the one from the representative (i.e., C_a^{480213}). Therefore, players $M+1, M+2, \ldots$ will announce the correct color. What about the first M players? Since the number M is known to players $1, \ldots, M-1$, we can apply the solution from the finite case with $M-1$ players. Again, the first player computes the corresponding sums of the observed colors from $2, 3, \ldots, M-1$, ignoring everything that comes after $M-1$. This way, players $2, 3, \ldots, M-1$ receive sufficient information to announce their hat colors correctly, as do the players $M+1, M+2, \ldots$. Player M's answer is certainly incorrect, and thus, the expected number of false answers is $2-1/m$. If this thought experiment feels a bit unsettling, we encourage you to take a walk outside. The world is beautifully finite.

C.3 CHAPTER 5: OPERA SINGERS AND INFORMATION THEORY

Uniform Distributions and Entropy

We will now prove that the uniform distribution indeed maximizes the entropy. Consider a random variable that takes the values $x_1, \ldots, x_m$ with probabilities $p_1, \ldots, p_m$, respectively. We now want to prove that $H(p_1, \ldots, p_m)$ is maximized for the uniform distribution (that is, if all of the p_i equal $1/m$).

The statement follows from a well-known mathematical inequality, called the *Gibb's inequality*. We start with the observation that for all $x > 0$, we have

$$\log_2 x \leq \frac{x-1}{\ln 2} \tag{C.3}$$

with equality if and only if $x = 1$. Here, we write $\ln 2 := \log_e 2$ to stress that the logarithm is taken to the base $e = \exp(1)$. There are several ways of seeing why equation (C.3) holds. It is probably easiest if you draw the two graphs. Regarding a formal proof, if you know what "strong convexity" means, you can also prove that $x \mapsto \ln x$ is strictly convex with $x \mapsto x - 1$ being the tangent at the graph at $x = 1$. From equation (C.3), it follows that, for any i, we have

$$p_i \log_2 \frac{1}{mp_i} \leq p_i \frac{\frac{1}{mp_i} - 1}{\ln 2}$$

with equality if and only if $mp_i = 1$. This even holds if we sum over all i. That is,

$$\sum_{i=1}^{m} p_i \log_2 \frac{1}{mp_i} \leq \sum_{i=1}^{m} p_i \frac{\frac{1}{mp_i} - 1}{\ln 2}$$

with equality if and only if for all i, $p_i = 1/m$. Transforming both sides yields

$$-\log_2 m + \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i} \leq \frac{1-1}{\ln 2} = 0$$

with equality if and only if for all i, $p_i = 1/m$. This proves that

$$H(p_1, \ldots, p_m) \leq \log_2 m$$

with equality if and only if for all i, $p_i = 1/m$. Thus, the entropy is maximized for the uniform distribution.

Stepwise Strategies Are Not Always Optimal

In section 5.4, we proposed to optimize the average or minimal information content step-wise, which is often called a "greedy" strategy. We now provide an example that shows that

such an approach does not necessarily yield an optimal strategy. Consider the set of hypotheses

$$\Omega = \{1, 2, 3, \ldots, 24\}.$$

Suppose that there are only 4 questions that we can ask:

$$O_1 : \begin{cases} 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8, & 9 & \mapsto & 0 \\ 10, & 11, & 12, & 13, & 14, & 15, & 16, & 17, & 18 & \mapsto & 1 \\ & & & 19, & 20, & 21, & 22, & 23, & 24 & \mapsto & 2 \end{cases}$$

$$O_2 : \begin{cases} & & 1, & 2, & 3, & 10, & 11, & 12, & 21 & \mapsto & 0 \\ 4, & 5, & 6, & 13, & 14, & 15, & 22, & 23, & 24 & \mapsto & 1 \\ & 7, & 8, & 9, & 16, & 17, & 18, & 19, & 20 & \mapsto & 2 \end{cases}$$

$$O_3 : \begin{cases} 1, & 4, & 7, & 10, & 13, & 16, & 19, & 21, & 22 & \mapsto & 0 \\ & 2, & 5, & 8, & 11, & 14, & 17, & 20, & 23 & \mapsto & 1 \\ & & 3, & 6, & 9, & 12, & 15, & 18, & 24 & \mapsto & 2 \end{cases}$$

$$O_4 : \begin{cases} & 1, & 4, & 6, & 7, & 8, & 9, & 12, & 16 & \mapsto & 0 \\ & 5, & 11, & 14, & 15, & 17, & 18, & 19, & 20 & \mapsto & 1 \\ & 2, & 3, & 10, & 13, & 21, & 22, & 23, & 24 & \mapsto & 2 \end{cases}$$

We have $\#\Omega = 24$, and the questions have 3 possible answers. We thus need at least 3 questions to find the secret number ω^*.

The strategy maximizing both the average and the minimal information content in each step would clearly tell us to start with the question O_4. All of its outcomes, $P(O_4 = 0) = P(O_4 = 1) = P(O_4 = 2) = 8/24 = 1/3$, have equal probability, and the average (or minimal) information equals $\log_2 3 \approx 1.585$ bits. Starting with O_4, however, will not allow us to always identify the secret ω^*. Suppose that the answer to O_4 equals $O_4(\omega^*) = 2$. If we then continue with O_1, and we receive the answer $O_1(\omega^*) = 2$, no matter whether we continue with O_2 or O_3 as a third question, there is a pair of two different ω values that we cannot distinguish between. This is shown by the first 4 rows of the following table:

ω^*	O_4	O_1	O_2	O_3
23	2	2	1	✗
24	2	2	1	✗
21	2	2	✗	0
22	2	2	✗	0
13	2	✗	1	0
22	2	✗	1	0

In fact, the table shows that in our example, there is no strategy starting with O_4 that is certain to identify ω^*. Even if we continue with O_2 rather than O_1 (and then use O_3 or O_1) or if we continue with O_3 (and then use O_1 or O_2), there are always two different ω^* values that yield the same answers.

If we start with O_1 instead of O_4, however, and use the questions O_1, O_2, O_3, we are guaranteed to always identify ω^*: There are no two ω^* values that map to the same 3-tuple.

This counterexample is constructed to show that we cannot expect optimality when performing a greedy search strategy. In most practical situations, however, the greedy approach might be a good way to start.

C.4 CHAPTER 6: ANIMAL MATCHING AND PROJECTIVE GEOMETRY

In the incidence matrix A, each row corresponds to a player, and each column corresponds to an animal. A value of 1 in position (k, j) of the incidence matrix (we write $A_{k,j} = 1$ in this case) indicates that player k has taken animal j.

In the remainder of this section, we want to relate the properties of A to the constraint in the game that every pair of players needs to hold at least 1 animal in common. We also make an argument for why all 7 players have to hold exactly 3 animals for a valid solution.

First, $A^\top$ is the so-called "transpose" of A. The transpose $A^\top$ can be thought of as A with interchanged rows and columns.

The entry (k, j) in matrix $A^\top$ is equal to the entry (j, k) in A. The dimensions of A and $A^\top$ are only the same if the number of rows and columns is equal (which is true here). An interesting matrix to study is

$$AA^\top,$$

the matrix multiplication of A with $A^\top$ (see appendix B.7 for a definition of a matrix multiplication). The matrix $AA^\top$ has the entries (using the definition of a transpose)

$$(AA^\top)_{k,j} = \sum_{\text{animal } \ell} A_{k,\ell}(A^\top)_{\ell,j} = \sum_{\ell} A_{k,\ell}A_{j,\ell}.$$

The entry (k, j) of matrix $AA^\top$ thus uses rows k and j of A, corresponding to players k and j, and is summing over all columns (animals) the product $A_{k,\ell}A_{j,\ell}$ of the corresponding entries of A. This product is usually 0 and can only be 1 if both $A_{k,\ell} = 1$ and $A_{j,\ell} = 1$ (that is, both players need to have animal ℓ in their possession). By summing over all animals, we get the total number of animals that both players have in common. If we want to guarantee that every pair of players has at least 1 animal in common, then we need to make sure that for all $k, \ell \in \{1, \ldots, 7\}$,

$$(AA^\top)_{k,\ell} \geq 1.$$

The diagonal part of $AA^\top$ will always be equal to the diagonal part of $3 \cdot \mathrm{Id}$ (that is, a matrix with 0s everywhere except for entries equal to 3 on the diagonal), as each row of A contains exactly three 1s. Assume that the incidence matrix satisfies

$$AA^\top = 2 \cdot \mathrm{Id} + J,$$

where Id is the identity matrix (1s on the diagonal and 0s everywhere else), and J is a matrix with all entries identically equal to 1. Then all pairs of players have 1 animal in common. Then the way that the audience picks the pairs does not matter, and the players will always win.

Is there another possible solution where, for example, some entries in $AA^\top$ take the value 2, so that some pairs of players

have 2 animals in common? To answer the question, it helps to look at the sum over all entries of $AA^\top$. In any arrangement where each player has 3 distinct animals, each of the 7 animals is present for 3 players. There will thus always be $3^2 = 9$ player-pairs (k, ℓ) that share the same animal, counting also cases where $k = \ell$ (diagonal entries of $AA^\top$) and counting both (k, ℓ) and (ℓ, k) if $k \neq \ell$ (that is, looking at upper-triangular and lower-triangular parts of $AA^\top$). Each animal thus contributes $3^2 = 9$ to the sum of all entries of $AA^\top$. Summing over all animals, we always get the value $7 \cdot 3^2 = 63$ for the sum of $AA^\top$ over all entries. This sum-constraint is satisfied, for example, by the solution $AA^\top = 2 \cdot \mathrm{Id} + J$. If we now find a solution that deviates from $AA^\top = 2 \cdot \mathrm{Id} + J$ by, for example, allowing some pairs to have 2 or more animals in common (which implies a value of 2 or more on the off-diagonal elements of $AA^\top$), then the sum constraint implies that this value has to be compensated for by some pairs not having any animal in common, with a corresponding entry of 0 on the off-diagonal elements. Hence $AA^\top = 2 \cdot \mathrm{Id} + J$ needs to be fulfilled for any solution that guarantees success.

The task is thus to create an incidence matrix A of dimension 7×7 that satisfies $AA^\top = 2 \cdot \mathrm{Id} + J$. In other words, the total number of animal matches is just enough to give each pair of players 1 animal to match. If any pair of players has more than 1 animal in common, then other pairs will have no animal in common.

C.5 CHAPTER 8: THE FALLEN PICTURE AND ALGEBRAIC TOPOLOGY

We now provide a formal introduction to the fundamental group. This concept is certainly not necessary for understanding chapter 8, but if you have some mathematical training, you might be wondering how to formalize the concepts introduced in chapter 8. In particular, if you know when functions $f : [0, 1] \to \mathbb{R}^n$ and $F : [0, 1] \times [0, 1] \to \mathbb{R}^n$ for some $n \in \mathbb{N} \setminus \{0\}$ are

said to be continuous, we hope that the following sequence of definitions will make sense.

Let X be a space that is a subset of $\mathbb{R}^n$ for some $n \in \mathbb{N} \setminus \{0\}$. We have the following definitions. A *loop around* $x_0 \in X$ (for simplicity, we will simply call this a "loop") is a continuous map

$$f : [0, 1] \to X$$

with $f(0) = f(1) = x_0$. A *homotopy of loops* is a collection of loops

$$f_t : [0, 1] \to X, \quad 0 \leq t \leq 1,$$

such that the map

$$F : \begin{array}{ccc} [0,1] \times [0,1] & \to & X \\ (s,t) & \mapsto & f_t(s) \end{array}$$

is continuous. Two loops g and h are *homotopic* if there is a homotopy $(f_t)_t$ with

$$f_0 = g \qquad \text{and} \qquad f_1 = h.$$

This defines an equivalence relation $\sim$ between loops. We identify loops with their equivalence classes. The concatenation of loops defines a group action:

$$g \circ h : \begin{array}{ccc} [0,1] & \to & X \\ t & \mapsto & \begin{cases} g(2t) & \text{if } t \leq 0.5 \\ h(2t-1) & \text{otherwise.} \end{cases} \end{array}$$

The *fundamental group* $\Pi_1(X, x_0)$ is defined as the set of all loops modulo the equivalence relation $\sim$ with concatenation as the group action. One can show that for path-connected spaces X, the fundamental group does not depend on x_0.