THE FUNDAMENTAL GROUP, COVERING SPACES AND TOPOLOGY IN BIOLOGY

KRZYSZTOF BARTOSZEK[1] AND JUSTYNA SIGNERSKA[2]

Abstract. We give a short introduction to homotopy theory. We pass to the concepts of a *pointed space* (X, x_0), the *fundamental group* of X, a *simply connected space* (with the example of the *space contractible to a point*), introduce basic concepts of *covering spaces* (e.g. *covering map/space, fiber over x, Path lifting Theorem*). With the use of the exponential map and the idea of the *index* of a loop, we show that the fundamental group of the circle S^1 is isomorphic to the integers $\mathbb{Z}$ with addition. We mention some other interesting fundamental groups (e.g. the fundamental group of a torus or of the *figure eight*). We also present some very interesting applications of topological concepts in Molecular Biology.

Algebraic topology tries to connect topological spaces with algebraical objects in such a way that topological problems can be translated into algebraical problems which can possibly be easier to solve. This paper is an introduction into the theory of homotopy and the basic concepts that concern it. Apart from the homotopy theory we present at the end of the paper interesting applications of topology in Molecular Biology. In the paper I denotes the closed interval $[0, 1]$.

1. Homotopy

Definition 1 (Homotopy)**.** *Let X and Y be topological spaces. Let $X' \subset X$ and $f_0, f_1 : X \to Y$ be continuous and agree on X'. f_0 is* homotopic *to f_1 relative to X' if there exists a continuous map $F : X \times I \to Y$ such that $F(x, 0) = f_0(x)$, $F(x, 1) = f_1(x)$ for $x \in X$ and $F(x, t) = f_0(x)$ for $x \in X'$ and $t \in I$.*

If f_0 and f_1 are homotopic relative to X' we write $f_0 \simeq f_1$ *rel* X'. If $X' = \emptyset$ we omit writing *rel* X'.

Example 1.

Let $X = Y = \mathbb{R}$ and $f_0(x) = x$, $f_1(x) = 0$ $x \in \mathbb{R}$. $f_0 \simeq f_1$ *rel* $\{0\}$ because the homotopy F is $F : \mathbb{R} \times I \to \mathbb{R}$ $F(x, t) = (1 - t)x$. We can see that F(x,0)=x, F(x,1)=0, F(0,t)=0.

Example 2.

Let $X = Y = I$ and $f_0(t) = t$, $f_1(t) = 0$ $t \in I$. $f_0 \simeq f_1$ *rel* $\{0\}$ because the homotopy F is $F : I \times I \to I$ $F(t, t') = (1 - t')t$. We can see that

[1] Student of the Department of Mathematics, Gdańsk University of Technology, ul. Narutowicza 11/12, 80-952 Gdańsk, Poland, *E-mail: kbart@sphere.pl* .

[2] Student of the Department of Mathematics, Gdańsk University of Technology, ul. Narutowicza 11/12, 80-952 Gdańsk, Poland, *E-mail: jussig@wp.pl* .

F(t,0)=t, F(t,1)=0,F(0,t')=0.

A homotopic map can be thought of as a continuous transformation of the function f_0 into f_1 on the set X' during the time interval I. An important type of homotopy is the homotopy of paths with the same ends about which more later will be written.

Theorem 1. *Homotopy relative to X' is an equivalence relation in the set of continuous maps $X \to Y$.*

PROOF Reflexivity: $f \simeq f$ because let $F(x,t) = f(x)$.
Symmetry: Let F be the homotopy for $f_0 \simeq f_1$ *rel* X'. Then the homotopy F' for $f_1 \simeq f_0$ *rel* X' is $F'(x,t) = F(x, 1-t)$.
Transitivity: Let F_1 be the homotopy for $f_0 \simeq f_1$ *rel* X' and F_2 be the homotopy for $f_1 \simeq f_2$ *rel* X'. Then F_3 defined

$$F_3(x,t) = \begin{cases} F_1(x,2t) & 0 \leq t \leq \frac{1}{2} \\ F_2(x,2t-1) & \frac{1}{2} \leq t \leq 1 \end{cases}$$

is the homotopy for $f_0 \simeq f_2$ *rel* X'. F_3 is continuous because both its restrictions are continuous and equal for $t = \frac{1}{2}$ (ie. $F_1(x,1) = f_1(x) = F_2(x,0)$).

Q.E.D.

Theorem 2. *Composites of homotopic maps are homotopic.*

Definition 2 (Pointed Space)**.** *We say that a topological space X is* pointed *if X is non empty and it has a base point $x_0 \in X$. We denote such a space by (X, x_0).*

Definition 3 (Contractible Space)**.** *A topological space X is said to be* contractible *if its identity map is homotopic to some constant map of X to itself ie. $f(x) = x$ and $f \simeq c$.*

From examples 1 and 2 we can see that $\mathbb{R}$ and I are contractible.

Theorem 3. *Any two maps of an arbitrary space to a contractible space are homotopic.*

PROOF Let Y be a contractible space and let its identity map be homotopic to the map c (constant in Y). Let $f_0, f_1 : X \to Y$ be arbitrary. By theorem 2 $f_0 \simeq cf_0$ and $f_1 \simeq cf_1$. Because $cf_0 = cf_1$ therefore by theorem 1 we have $f_0 \simeq f_1$.

Q.E.D

2. The Fundamental Group

An interesting case of homotopy is the homotopy of paths (ie. maps $I \to X$).

Definition 4. *Let σ and τ be two paths in X with the same end points ($\sigma(0) = \tau(0) = x_0$ and $\sigma(1) = \tau(1) = x_1$). We say that σ and τ are* homotopic relative to the end points *($\sigma \simeq \tau$ rel $\{0,1\}$) if there exists a map $F : I \times I \to X$ such that $F(t,0) = \sigma(t), F(t,1) = \tau(t), F(0,t') = x_0, F(1,t') = x_1$ $t, t' \in I$*

We can see that the above definition is the general definition of homotopy applied to paths and the set $\{0,1\} \subset I$. Therefore the homotopy of paths will be an equivalence relation in the set of paths $I \to X$ for a given space X. Also we can think as before of homotopy as a continuous transformation of one path to another during the time interval I.

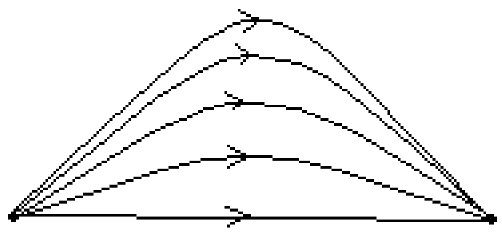

FIGURE 1. Path transformation.

We define the multiplication of two paths as

$$\sigma\tau(t) = \begin{cases} \sigma(2t) & 0 \le t \le \frac{1}{2} \\ \tau(2t-1) & \frac{1}{2} \le t \le 1 \end{cases}$$

Because homotopy is an equivalence relation we can speak of multiplying (as above) the equivalence class of one path with the equivalence class of another path (presuming their ends are the same).

Theorem 4. *Let $\pi_1(X, x_0)$ be the set of classes of homotopy of loops in X at the point x_0. With the above definition of multiplication $\pi_1(X, x_0)$ is a group. The identity element being the constant loop in x_0 and the inverse of $[\sigma]$ being $[\sigma^{-1}]$ defined $\sigma^{-1}(t) = \sigma(1-t), 0 \le t \le 1$.*

Theorem 5. *If X is a path-connected space then for all points $x_0 \in X$ the groups $\pi_1(X, x_0)$ are isomorphic.*

In the above case (X is path connected) we call $\pi_1(X, x_0)$ the fundamental group of X and write it as $\pi_1(X)$.
We can now look at contractible spaces as being a special case of a more general class of spaces.

Definition 5 (Simply Connected Space)**.** *A topological space (X, A) is said to be a* simply connected space *if it is path connected and the fundamental group of X is trivial.*

A simply connected space is therefore a path connected space in which all paths between two points can be continuously transformed into each other. Examples of simply connected spaces are $\mathbb{R}^2$ and S^n $(n \ge 2)$.

Theorem 6. *A contractible space is a simply connected space.*

A concept of the fundamental group was introduced by French mathematician, Henri Poincaré (1854-1912) in his *Analysis situs* in 1895. Poincaré is said to be the originator of algebraic topology. One of his famous quotations states that: *Mathematics is the art of giving the same name to different things* (as opposed to the quotation *Poetry is the art of giving different names to the same thing*). In fact, the fundamental group (and higher homotopy groups) can be very useful in classifying spaces.

Let (X, b) and (Y, c) be pointed topological spaces. A map $f : (Y, c) \to (X, b)$ is a continuous function f that satisfies $f(c) = b$. Then the map $f_* : \pi_1(Y, c) \to \pi_1(X, b)$ is defined by making the path γ in Y correspond to the path $f \circ \gamma$ in X. This correspondence respects homotopy classes.

Theorem 7. *If $f : Y \to X$ is a homeomorphism and if $c \in Y$ and $b = f(c)$, then f_* is an isomorphisms of $\pi_1(Y, c)$ and $\pi_1(X, b)$*

Thus two spaces could be shown to be nonhomeomorphic by showing that their fundamental groups were not isomorphic. However, given the topological space X, we need to be able to compute its. And here we sometimes need the concept of *covering maps* and *covering spaces*.

3. Covering spaces

Let X and E be topological spaces. Let $p : E \to X$ be a continuous map.

Definition 6 (Evenly covered set)**.** *We say that an open subset U of X is* evenly covered *by p if the inverse image $p^{-1}(U)$ is a union of disjoint open subsets of E, each of which is mapped homeomorphically by p onto U.*

Definition 7 (Covering map)**.** *The map p is a* covering map *if p continuously maps E onto X such that each $x \in X$ has an open neighborhood which is evenly covered by p.*

In the above case, E is called a *covering space* over X. A covering map is also simply called a *cover*.

The so-called open cover, well known from elementary topology, is a special case of a covering space, when E is a disjoint union of a collection of open sets X_i with union X.

If $x \in X$, the set $p^{-1}(x)$ is called the *fiber* over x. From the definition of a covering map it follows that it is a discrete subspace of E and each $x \in X$ has an open neighborhood U such that $p^{-1}(U)$ is homeomorphic to $p^{-1}(x) \times U$. The subsets of $p^{-1}(U)$ mapped homeomorphically onto U are called the *sheets* of $p^{-1}(U)$. Of course, if U is connected, then the sheets of $p^{-1}(U)$ coincide with the connected components of $p^{-1}(U)$ (since connectedness is preserved by homeomorphism). Every cover $p : E \to X$ is a local homeomorphism which implies that X and its covering space E always share the same local topological properties. However, if X is *simply connected*, topological properties of X and E are held globally as well, with p being a homeomorphism of these two spaces.

A couple of elementary examples of covering maps can be provided.

Example 3.

Consider the complex plane with the origin removed $\mathbb{C}\backslash\{0\}$ and let n be a nonzero integer. Define the map $p : \mathbb{C}\backslash\{0\} \to \mathbb{C}\backslash\{0\}$ by $p(z) = z^n$. p is a covering map where every fiber $p^{-1}(x)$ has n elements.

Example 4.

The exponential map $p : \mathbb{R} \to S^1$ given by the formula:

$$p(t) = e^{2\pi i t}, \qquad t \in \mathbb{R} \tag{1}$$

is also a covering map. Indeed, let $x_0 = e^{2\pi i t_0}$ be a point in S^1. Choose $0 < \epsilon < \frac{1}{2}$ and let $U = \{e^{2\pi i t} : |t - t_0| < \epsilon\}$. Then $p^{-1}(U)$ is the disjoint union of the intervals $(m + t_0 - \epsilon, m + t_0 + \epsilon)$, $m \in \mathbb{Z}$, each of which is mapped homeomorphically by p onto U. Later we will use this example of a covering map to derive the fundamental group of the unit circle S^1.

Example 5.

Consider the n-dimensional projective space P^n, i.e. the quotient space obtained from the unit sphere S^n by identifying two antipodal points. Let $p : S^n \to P^n$ be the quotient map. This a covering map. If $x_0 \in S^n$, $\epsilon > 0$ is efficiently small and $V = \{x \in S^n : |x - x_0| < \epsilon\}$, then $U = p(V)$ is an open subset of P^n and $p^{-1}(U)$ is the disjoint union of V and $-V$, where both V and $-V$ are mapped homeomorphically by p onto U. Here every fiber consists of two points.

Theorem 8. *Products of covering spaces are covering spaces.*

Example 6.

$\mathbb{R}^n$ is the covering space over the n-dimensional torus $T^n = S^1 \times ... \times S^1$ with the covering map $p : \mathbb{R}^n \to T^n$ given by:

$$p(x_1, ..., ..., x_2) = (e^{2\pi i x_1}, ..., ..., e^{2\pi i x_n}).$$

We will now pass to quite significant theorems in topology involving covering maps. Suppose that E, X and Y are topological spaces and $p : E \to X$ is a covering map. Let $f : Y \to X$ be a map. It turns out that often we need to determine whether there exists a map $g : Y \to E$ such that $p \circ g = f$. In this case, g is called a *lift* of f. The situation is represented by a diagram:

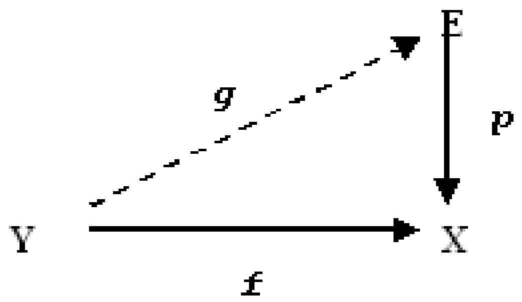

FIGURE 2. A diagram of the lift g.

When $p \circ g = f$, the diagram *commutes*. The following lemma tells about the uniqueness of lifts:

Lemma 1. *Let $p : E \to X$ be a covering map and let Y be a connected topological space. Let $f : Y \to X$ be a map and let $g, h : Y \to E$ be two lifts of f. If $g(y) = h(y)$ for some point $y \in Y$, then $g = h$.*

The next theorem will be of high importance for us:

Theorem 9 (Path Lifting Theorem)**.** *Let $p : E \to X$ be a covering map, let $\gamma : I \to X$ be a path, and let e_0 satisfy $p(e_0) = \gamma(0)$. Then there exist a unique path $\alpha : I \to E$ such that $\alpha(0) = e_0$ and $p \circ \alpha = \gamma$.*

PROOF. For each $x \in X$ we choose an open neighborhood $U(x)$ that is evenly covered by p. Then the open sets $\gamma^{-1}(U(x))$, $x \in X$, form an open cover of I.Since I is compact, we can find $0 = s_0 < s_1 ... < s_m = 1$ and sets $U_1, U_2 ... U_m$ such that:

$$\gamma([s_{j-1}, s_j]) \subset U_j, \quad 1 \le j \le m.$$

Since $p(e_0) = \gamma(0) \in U_1$, there is an open neighborhood V_1 of e_0 that is mapped homeomorphically by p onto U_1. We define a function α on the interval I as follows:

$$\alpha = (p\,|_{V_1})^{-1} \circ \gamma \quad \text{on} \quad [0, s_1].$$

Then $\alpha(s)$ is the unique point of V_1 covering $\gamma(s)$, $\alpha(0) = e_0$ and $p \circ \alpha = \gamma$ on $[0, s_1]$. We repeat the procedure with $e_0 = \alpha(0)$ replaced by $\alpha(s_1)$ and U_1 replaced by U_2 extending α to the interval $[s_1, s_2]$. After m steps the whole γ path is lifted.

Uniqueness follows from the previous lemma.

Q.E.D.

If (E, e_0) and (X, x_0) are two pointed spaces, then $p : (E, e_0) \to (X, x_0)$ is a covering map when $p : E \to X$ is a covering map satisfying $p(e_0) = x_0$.

Actually, we can lift not only a single path but also a whole family of paths depending continuously on a parameter in such a way that the lifted paths also depend continuously on the parameter. This follows from:

Theorem 10 (Homotopy Lifting Theorem)**.** *Let $p : (E, e_0) \to (X, x_0)$ be a covering map. Suppose that a map $f : (Y, y_0) \to (X, x_0)$ has a lift $f' : (Y, y_0) \to (E, e_0)$. Then every homotopy $F : Y \times I \to X$ such that $F(y, 0) = f(y)$ for $y \in Y$ can be lifted to homotopy $F' : Y \times I \to E$ with $F'(y, 0) = f'(y)$ for $y \in Y$.*

A following immediate conclusion can be drawn from the above theorem:

Corollary 1. *If σ and τ are two paths in X starting at x_0 and $\sigma \simeq \tau$ rel $\{0, 1\}$, then $\sigma'_{e_0} \simeq \tau'_{e_0}$ rel $\{0, 1\}$. In particular, the lifts σ'_{e_0} and τ'_{e_0} end at the same point in E.*

The above statement allows us to define a function

$$\Phi : \pi_1(X, x_0) \to p^{-1}(x_0) \tag{2}$$

so that $\Phi([\gamma])$ is the terminal point of the lift of γ to E that starts at e_0 (γ - loop in X based at x_0).

Φ has the following important property:

Theorem 11 (Cardinality of fibers)**.** *Let $p : (E, e_0) \to (X, x_0)$ be a covering map and suppose that E is simply connected. Then Φ is a one-to-one correspondence of $\pi_1(X, x_0)$ and the fiber $p^{-1}(x_0)$.*

Let us now return to the covering map $p : (\mathbb{R}, 0) \to (S^1, 1)$, given by (1). Since $p^{-1}(1)$ coincides with the subset $\mathbb{Z} \subset \mathbb{R}$, the elements of the fundamental group $\pi_1(S^1, 1)$ are in one-to-one correspondence with the integers. In fact, this correspondence is a group isomorphism (we consider the group $(\mathbb{Z}, +)$). Let us look in detail how a loop in S^1 determines an integer.

Let γ be a loop in S^1 based at 1. The lift of γ is a map

$$h : I \to \mathbb{R}$$

satisfying:

$$e^{2\pi i h(s)} = \gamma(s), \quad 0 \le s \le 1 \tag{3}$$
$$h(0) = 0.$$

The terminal point $h(1)$ of h must be then an integer. This integer we will call the *index of* γ and denote it by $ind(\gamma)$ (it is the element of $p^{-1}(1)$ associated with $[\gamma]$ by (2)). In fact the following is true:

Theorem 12. *Two loops in S^1 based at 1 are in the same homotopy class if and only if they have the same index. The correspondence*

$$[\alpha] \to ind(\alpha)$$

is an isomorphism of $\pi_1(S^1, 1)$ and the integers $(\mathbb{Z}, +)$.

PROOF. The only thing which remains not proven for the above theorem is that this correspondence is a group homomorphism. It suffices to show that:

$$ind(\alpha_1\alpha_2) = ind(\alpha_1) + ind(\alpha_2)$$

where are α_1 and α_2 are arbitrary loops in S^1 based at 1.

Let us choose maps: $h_1, h_2 : I \to \mathbb{R}$ such that $h_1(0) = h_2(0) = 0$ and

$$\alpha_j(s) = e^{2\pi i h_j(s)}, \quad 0 \leq s \leq 1;\ j = 1, 2.$$

We define

$$h(s) = \begin{cases} h_1(2s), & 0 \leq s \leq 1/2; \\ h_1(1) + h_2(2s-1), & 1/2 \leq s \leq 1. \end{cases}$$

Then $h : I \to \mathbb{R}$ is continuous, $h(0) = 0$ and

$$(\alpha_1\alpha_2)(s) = e^{2\pi i h(s)}, \quad 0 \leq s \leq 1.$$

Hence:

$$ind(\alpha_1\alpha_2) = h_1(1) + h_2(1) = ind(\alpha_1) + ind(\alpha_2).$$

Q.E.D

The above elaboration shows that when given a loop γ in S^1, its homotopy class $[\gamma]$ is determined by the number of times this loop winds around the circle S^1 (if this number is negative, we wind the circle in the opposite direction to the given one).

Example 7.

Let us now again consider *real projective space* of dimension $n \geq 2$, P^n and the covering map $p : S^n \to P^n$ that was discussed earlier. Let e_0 be the north pole of S^n and set $x_0 = p(e_0)$. It can be checked that for $n \geq 2$, S^n is simply-connected. Hence we apply theorem 10 and conclude that $\pi_1(P^n, x_0)$ has exactly two elements (one element is the identity and the other is the homotopy class of of $p \circ \alpha$, where α is any path on S^n from the north pole to the south pole). Since any group with two elements is isomorphic to $\mathbb{Z}_2$:

$$\pi_1(P^n) \cong \mathbb{Z}_2, \quad n \geq 2.$$

The next theorem might also be very useful:

Theorem 13. *For given X, Y, x_0, y_0 the following is true:*

$$\pi_1(X \times Y, (x_0, y_0)) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0).$$

Example 8.

One of the immediate conclusions from this theorem is the fact that the fundamental group of a torus is $\mathbb{Z} \times \mathbb{Z}$ (since torus is homeomorphic to $S^1 \times S^1$).

Example 9.

Now let X be the figure eight space. Its fundamental group is quite interesting.

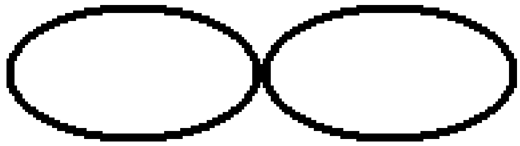

FIGURE 3. The figure eight.

It can be shown that every element of $\pi_1(X)$ can be expressed uniquely as a finite product

$$[\alpha]^{m_1}[\beta]^{m_2}[\alpha]^{m_3}...,$$

where m_1, m_2, ... are integers and $m_j \neq 0$ for $j \geq 2$ (only $m_1 = 0$ is allowed, otherwise the representation would not be unique). We may assume, to better see this, that α denotes a path in the figure eight along its right circle in the counterclockwise direction and β denotes a path along the left circle in the counterclockwise direction. Hence we conjecture that $\pi_1(X)$ is *the free group with two generators* $[\alpha]$ and $[\beta]$, namely it is (isomorphic to) $\mathbb{Z} * \mathbb{Z}$ (a formal proof of this fact would require more elaborate discussion). A group is called a free group if no relation exists between its group generators other than the relationship between an element and its inverse.

Theorem 14. *If $p : (E, e_0) \to (X, x_0)$ and $p' : (E', e_0') \to (X, x_0)$ are two covers of X and both covering spaces E and E' are simply connected, then there exists exactly one homeomorphism $\phi : (E', e_0') \to (E, e_0)$ such that $p \circ \phi = p'$.*

Definition 8 (Equivalent covers)**.** *If there exist a homeomorphism $\phi : (E', e_0') \to (E, e_0)$ such that $p \circ \phi = p'$, the two covers, p and p', are* equivalent.

If ϕ is a covering map rather than a homeomorphism, we say that p *dominates* p'.

Definition 9 (Universal cover)**.** *If (X, x_0) has a cover $(\widetilde{X}, \widetilde{x}_0) \to (X, x_0)$ such that the space $\widetilde{X}$ is simply connected, then this cover is called* universal cover *of (X, x_0).*

Theorem 14 states that the universal cover is unique up to equivalence.

For a topological space X, a universal cover might not exist. Since spaces X and $\widetilde{X}$ are locally homeomorphic, "small" loops in X should be homotopic *rel* $\{0, 1\}$ to a trivial loop. Hence the necessary condition, for a topological space X to have a universal cover, is *semi-locally simply connectedness*:

Definition 10 (Semi-locally simply connected space)**.** *If each point in $x \in X$ has an open neighborhood U such that each loop in U based at x is homotopic rel $\{0, 1\}$ to a constant loop based at x, then X is called* semi-locally simply connected

The prefix semi- refers to the fact that the homotopy which takes the loop to the trivial loop can leave U and travel to other parts of X.

There is also a sufficient condition for a universal cover to exist.

Theorem 15. *The space X has a universal cover if and only if it is path-connected, locally path connected and semi-locally simply connected.*

Corollary 2. *Every connected manifold has a universal cover (with a covering space being other manifold).*

Example 10.

Let C_n denote a circle on a plane with the center $(\frac{1}{n}, 0)$ and the radius equal to $\frac{1}{n}$. Let

$$X = \bigcup_n C_n.$$

Then X does not have a universal cover since it is not semi-locally simply connected (the condition is violated at $(0,0)$).

The covering space $\widetilde{X}$ of X of the universal cover can be constructed as a certain space of paths in X: fix $x_0 \in X$ and denote $\widetilde{X}$ as the set of all pairs $(x, [\gamma])$, where $x \in X$ and γ is a path in X from x_0 to x.

Definition 11 (deck transformation)**.** *A* deck transformation *(or* covering transformation*) of a cover $p : E \to X$ is a homeomorphism $\phi : E \to E$ such that $p \circ \phi = p$.*

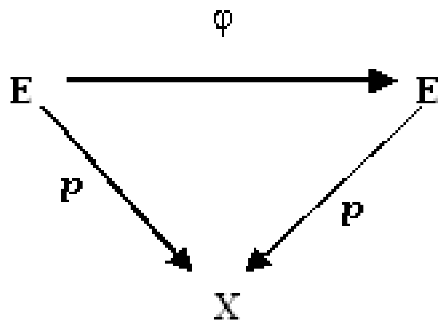

Figure 4. Deck transformation.

The set of all deck transformations of p forms a group under composition, the *deck transformation group $Aut(p)$*. A deck transformation permutes the elements of each fiber.

Theorem 16. *A deck transformation group of a universal cover $p : E \to X$ is isomorphic to the fundamental group $\pi_1(X)$.*

In fact, if we are given simply connected covering space E of X and its deck transformation group G, we not only know that $\pi_1(X) \cong G$ but we can also "reconstruct" the space X itself (up to homeomorphism) as E/G.

One of the strengths of algebraic topology is its wide degree of applicability to other fields. Nowadays that includes fields like physics, differential geometry, algebraic geometry, number theory and other branches of science that one would sometimes not expect. Here we present some examples of applications in biology.

4. Topology in Biology

The authors of [5] have defined an interesting method of classifying RNA structures using the concept of the *genus*. This concept will be presented here. Each RNA structure is represented by a diagram (figure 5). The backbone of the RNA (the main chain) can be open or closed depending how it will be more convenient. These diagrams are called *double line diagrams*. The genus g is defined as follows (when the backbone is open)

$$g = \frac{P - L}{2} \tag{4}$$

where P is the number of double lines (pairings) and L is the number of closed loops made by the single lines. When we close the backbone g is equal to the amount of "handles" a sphere must have so that the diagram can drawn on it without any crossings. The genus is a topological invariant of the diagram. We can see that a

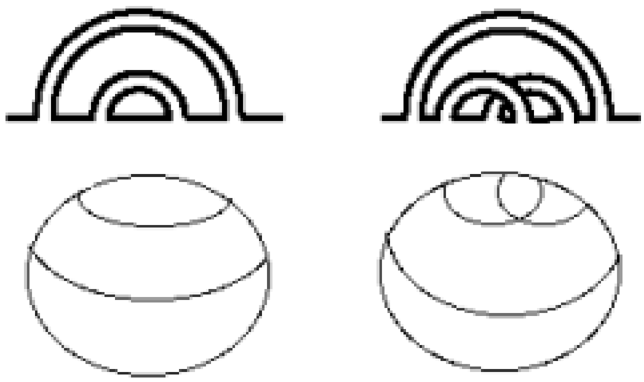

Figure 5. Left diagram of genus 0 right of genus 1. In the bottom the same diagrams with a closed RNA backbone [5].

figure of genus 0 is a planar figure, genus 1 implies that it can be drawn on a torus. In order to use these concepts two definitions have to be introduced.

Definition 12 (Irreducibility)**.** *A diagram is said to be* irreducible *if it cannot be broken down into two disconnected parts by a single cut* (figure 6).

For example if a diagram is irreducible a cut on the backbone is not enough to disconnect it.

Figure 6. A reducible diagram [5].

Definition 13 (Nesting). *A diagram is said to be* nested *in another if it can be removed by cutting in two places while the rest stays connected* (figure 7).

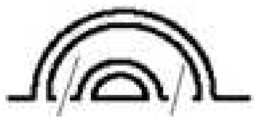

FIGURE 7. A nested diagram [5].

The genus of a reducible diagram is the sum of the genera of its irreducible components and also its genus is the sum of the genera of its nested components. A pseudoknot (a RNA secondary-structure) is said to be *primitive* if it is non-nested and irreducible. It has been shown in [11] that there are 8 irreducible RNA pseudoknots of genus 1. Of these 4 are very common, 2 are rarely seen and 2 have never been met. The authors of [5] classified two databases, a pseudoknot base (*Pseudobase*) and the world wide Protein Data Bank (*wwPDB*). 96, 7% of all pseudoknots were of genus 1 with the same topology. In the classification of proteins an interesting fact was observed. When one compares the genus with the length of the RNA the genus is much lower then one would expect of a genus of a random sequence. This suggests that there is some design in the structure of the RNA and that some information is carried by the shape itself. At the moment there is intensive research done to discover what information is carried by the shape of RNA and other molecular structures.

Another problem in Molecular Biology is the classification of proteins. Often identical proteins have different 3D-structures. Often topological concepts are introduced to aid. In [6] the concept of homotopy was used to compare two proteins. Two proteins were considered similar in structure if there is a rigid body transformation that places one protein on the other and it allows for atom alignments.

REFERENCES

[1] About algebraic topology. http://www.math.rochester.edu/people/faculty/jnei/algtop.html.
[2] Famous quotations by Poincare. http://www-groups.dcs.st-and.ac.uk/ h̃istory/Quotations/Poincare.html.
[3] History of topology. http://www-groups.dcs.st-and.ac.uk/h̃istory/HistTopics/Topology_in_mathematics.html.
[4] Wikipedia: Simply connected space. http://en.wikipedia.org/wiki/Simply_connected.
[5] M. Bon, G. Vernizzi, H. Orland, and A. Zee. Topological classification of RNA structures, 2006. http://arxiv.org/pdf/q-bio.BM/0607032.
[6] M. Erdmann. Protein similarity from knot theory and geometric convolution. Technical report, School of Computer Science Carnegie Mellon University, 2003.
[7] T. W. Gamelin and R. Everist Greene. *Introduction to Topology.* DOVER PUBLICATIONS, New York, 1999.
[8] M. J. Greenberg. *Wykłady z topologii algebraicznej.* PWN, Warszawa, 1980.
[9] D. Higgins and W. Taylor, editors. *Bioinformatics Sequence, structure and databanks.* Oxford University Press, 2003.
[10] K. Janich. *Topologia.* PWN, Warszawa, 1996.
[11] M. Pillsbury, H. Orland, and A. Zee. Steepest descent calculation of RNA pseudoknots. *Physical Review E*, 72, 2005.
[12] E. H. Spanier. *Algebraic Topology.* McGraw-Hill, 1966.