A HOL Light formalization of singular homology

John Harrison

Amazon Web Services

Hales60, 21st June 2018 (09:00-09:50)

Tom Hales is famous for many things . . .

. . . including the first Jordan Curve Theorem formalization

The first formal proof of the Jordan curve theorem was created by Hales (2007a) in the HOL Light system, in January 2005, and contained about 60,000 lines. Another rigorous 6,500-line formal proof was produced in 2005 by an international team of mathematicians using the Mizar system. Both the Mizar and the HOL Light proof rely on libraries of previously proved theorems, so these two sizes are not comparable. Nobuyuki Sakamoto and Keita Yokoyama (2007) showed that the Jordan curve theorem is equivalent in proof-theoretic strength to the weak König's lemma.

`https://en.wikipedia.org/wiki/Jordan_curve_theorem`

. . . and the rehabilitation of the original proof

However, Thomas C. Hales wrote:

> Nearly every modern citation that I have found agrees that the first correct proof is due to Veblen... In view of the heavy criticism of Jordan's proof, I was surprised when I sat down to read his proof to find nothing objectionable about it. Since then, I have contacted a number of the authors who have criticized Jordan, and each case the author has admitted to having no direct knowledge of an error in Jordan's proof.[4]

`https://en.wikipedia.org/wiki/Jordan_curve_theorem`

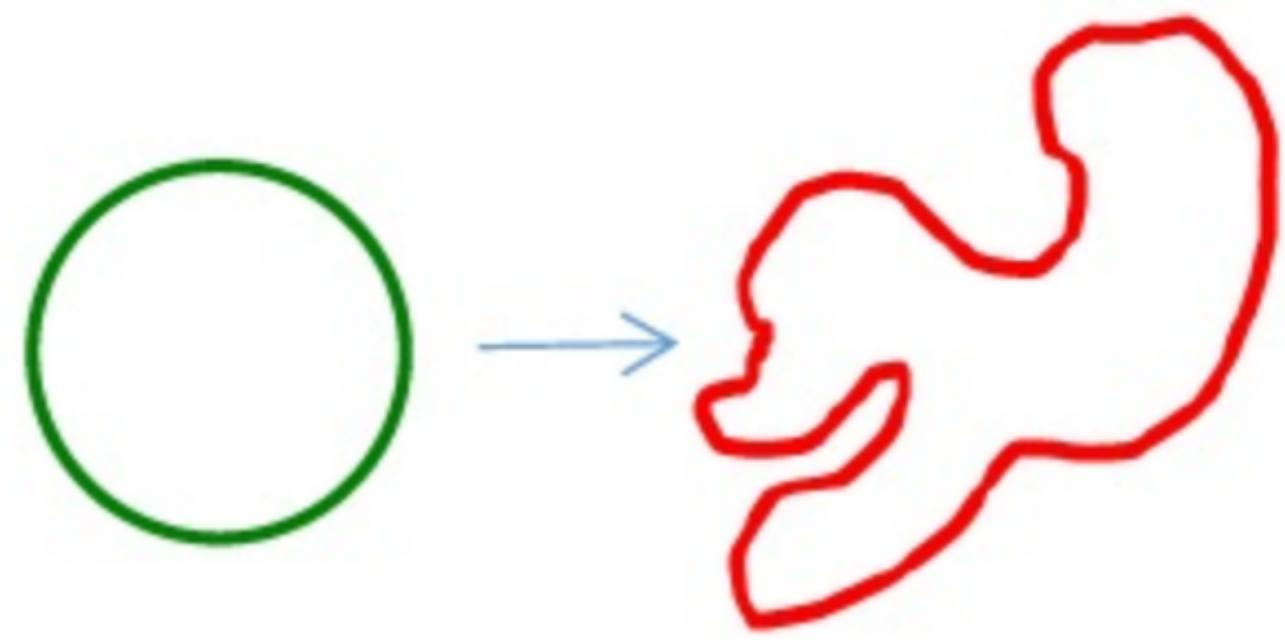
What is the Jordan Curve Theorem?

Every simple closed curve in the plane separates the plane into exactly two connected subsets, one bounded (the "inside") and one unbounded (the "outside").

- Curve: continuous function $\gamma : [0, 1] \to \mathbb{R}^2$
- Closed: $\gamma(0) = \gamma(1)$
- Simple: $\forall x\ y.\ \gamma(x) = \gamma(y) \Rightarrow x = y \vee \{x, y\} = \{0, 1\}$

Alternative point of view

Alternatively: a simple closed curve is a homeomorphic map $\gamma : S^1 \to \mathbb{R}^2$ out of the unit circle $S^1 = \{x \in \mathbb{R}^2 \mid |x| = 1\}$



Every subset of $\mathbb{R}^2$ homeomorphic to S^1 separates the plane into exactly two connected subsets, one bounded (the "inside") and one unbounded (the "outside").

The Jordan Curve Theorem in HOL Light

```
|- !c. simple_path c /\ pathfinish c = pathstart c
     ==> ?ins out.
           ~(ins = {}) /\ open ins /\ connected ins /\
           ~(out = {}) /\ open out /\ connected out /\
           bounded ins /\ ~bounded out /\
           ins INTER out = {} /\
           ins UNION out =
             (:real^2) DIFF path_image c /\
           frontier ins = path_image c /\
           frontier out = path_image c
```

The Jordan-Schoenflies Theorem in HOL Light

```
|- !g h f f'.
      simple_path g /\ simple_path h /\
      homeomorphism (path_image g,path_image h) (f,f')
      ==> ?k k'.
            homeomorphism ((:real^2),(:real^2)) (k,k') /\
            (!x. x IN path_image g ==> k x = f x) /\
            (!y. y IN path_image h ==> k' y = f' y) /\
            IMAGE k (path_image g) = path_image h /\
            IMAGE k (inside(path_image g)) =
              inside(path_image h) /\
            IMAGE k (outside(path_image g)) =
              outside(path_image h)
```

Proof uses not just topology but also complex analysis (Riemann Mapping Theorem, . . .)

Why isn't the Jordan Curve Theorem obvious?

- Continuous curves can have counterintuitive properties, even filling a "solid" shape (Peano's space-filling curves, Hahn-Mazurkiewicz theorem)

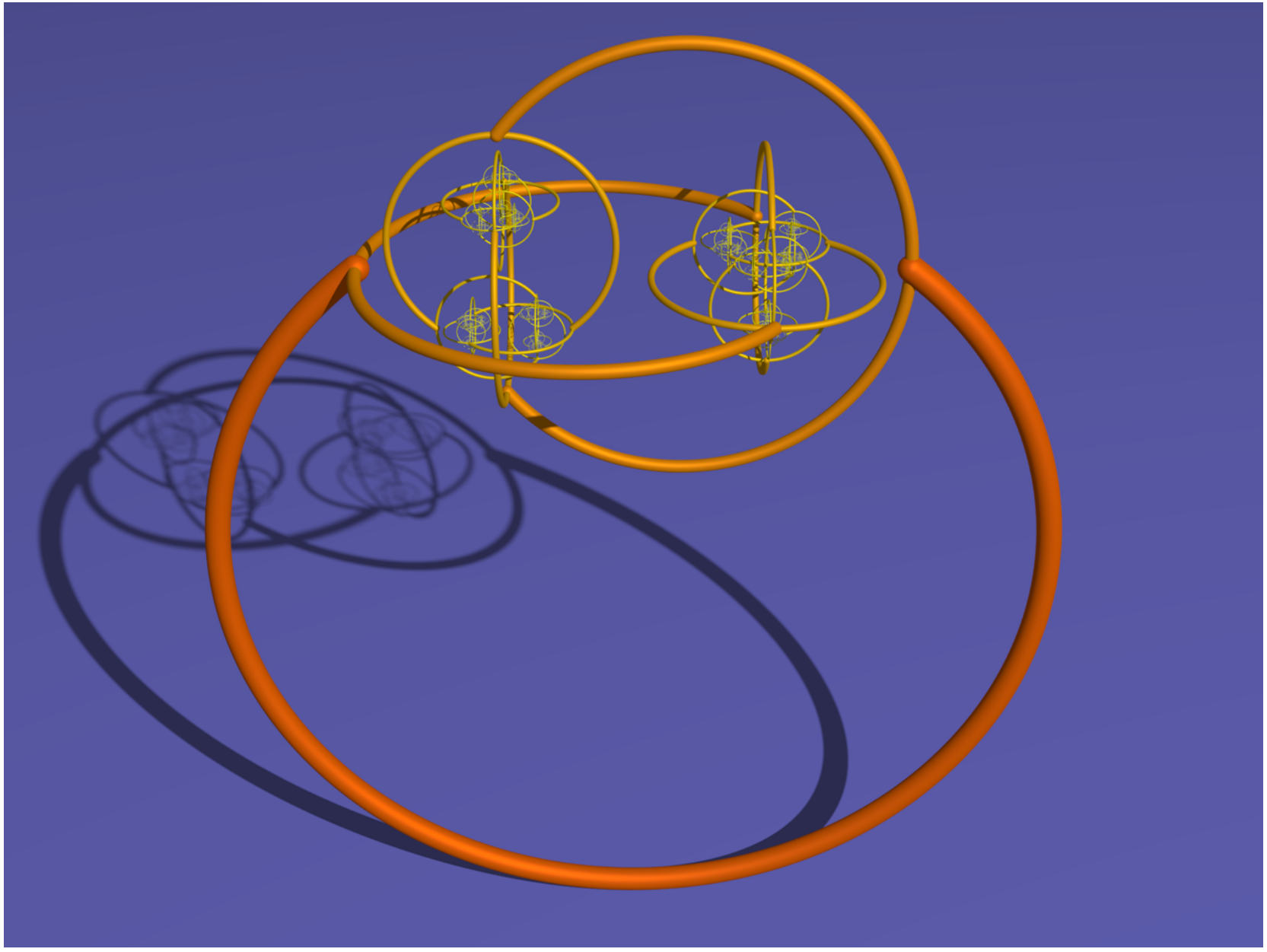
Why isn't the Jordan Curve Theorem obvious?

- Continuous curves can have counterintuitive properties, even filling a "solid" shape (Peano's space-filling curves, Hahn-Mazurkiewicz theorem)
- While *simple* curves can't fill 2-D regions, they can still have 2-D Lebesgue density $1 - \epsilon$ (Osgood, Knopp).

Why isn't the Jordan Curve Theorem obvious?

- Continuous curves can have counterintuitive properties, even filling a "solid" shape (Peano's space-filling curves, Hahn-Mazurkiewicz theorem)
- While *simple* curves can't fill 2-D regions, they can still have 2-D Lebesgue density $1 - \epsilon$ (Osgood, Knopp).
- Higher-dimensional analogs in $\mathbb{R}^N$ look equally plausible. But while Jordan does extend, Jordan-Schoenflies does not (Alexander).

Alexander's Horned Sphere



HOL Light Multivariate library

Partly as a result of Flyspeck, HOL Light is particularly strong in the area of topology, analysis and geometry in Euclidean space $\mathbb{R}^n$.

File	Lines	Contents
misc.ml	2540	Background stuff
metric.ml	23331	Metric spaces and general topology
homology.ml	**10539**	**Singular homology**
vectors.ml	10799	Basic vectors, linear algebra
determinants.ml	4776	Determinant and trace
topology.ml	36850	Topology of euclidean space
convex.ml	17848	Convex sets and functions
paths.ml	28185	Paths, simple connectedness etc.
polytope.ml	8940	Faces, polytopes, polyhedra etc.
degree.ml	9720	Degree theory, retracts etc.
derivatives.ml	5758	Derivatives
clifford.ml	979	Geometric (Clifford) algebra
integration.ml	26145	Integration
measure.ml	29998	Lebesgue measure

Multivariate theories continued

From this foundation complex analysis is developed and used to derive convenient theorems for $\mathbb{R}$ as well as more topological results.

File	Lines	Contents
complexes.ml	2249	Complex numbers
canal.ml	4019	Complex analysis
transcendentals.ml	7559	Real & complex transcendentals
realanalysis.ml	17556	Some analytical stuff on $\mathbb{R}$
moretop.ml	9583	Further topological results
cauchy.ml	23774	Complex line integrals

Credits: JRH, Marco Maggesi, Valentina Bruno, Graziano Gentili, Gianni Ciolli, Lars Schewe, ...

The awkward squad

There are some classic theorems saying something quite concrete about Euclidean space that seem difficult to prove ‘with bare hands’, though I tried . . .

The awkward squad

There are some classic theorems saying something quite concrete about Euclidean space that seem difficult to prove 'with bare hands', though I tried ...

- ☺ Brouwer's fixed-point theorem: Kuhn's explicit combinatorial proof with cubical subdivision

The awkward squad

There are some classic theorems saying something quite concrete about Euclidean space that seem difficult to prove 'with bare hands', though I tried ...

- ☺ Brouwer's fixed-point theorem: Kuhn's explicit combinatorial proof with cubical subdivision
- ☺ The hairy ball theorem: ad-hoc partial notion of topological degree following Dugundji

The awkward squad

There are some classic theorems saying something quite concrete about Euclidean space that seem difficult to prove 'with bare hands', though I tried ...

- ☺ Brouwer's fixed-point theorem: Kuhn's explicit combinatorial proof with cubical subdivision
- ☺ The hairy ball theorem: ad-hoc partial notion of topological degree following Dugundji
- ☹ The Jordan-Brouwer separation theorem: homotopic characterization following Borsuk a partial solution

The awkward squad

There are some classic theorems saying something quite concrete about Euclidean space that seem difficult to prove 'with bare hands', though I tried ...

- ☺ Brouwer's fixed-point theorem: Kuhn's explicit combinatorial proof with cubical subdivision
- ☺ The hairy ball theorem: ad-hoc partial notion of topological degree following Dugundji
- ☹ The Jordan-Brouwer separation theorem: homotopic characterization following Borsuk a partial solution
- ☹ The Borsuk-Ulam theorem: ???

The Jordan-Brouwer Separation Theorem in HOL Light

Thanks to the singular homology theory, we can now derive this general form of the Jordan Curve Theorem:

```
|- !s. 2 <= dimindex(:N) /\
      s homeomorphic sphere(vec 0,&1)
      ==> ?ins out.
            ~(ins = {}) /\ open ins /\ connected ins /\
            ~(out = {}) /\ open out /\ connected out /\
            bounded ins /\ ~bounded out /\
            ins INTER out = {} /\
            ins UNION out = (:real^N) DIFF s /\
            frontier ins = s /\
            frontier out = s
```

Basic idea of homology theory

For each $p \in \mathbb{Z}$, define a 'pth homology group' functor mapping topological spaces to groups

- Topological space X is mapped to its homology group $H_p(X)$
- A continuous map $f : X \to Y$ between topological spaces gives rise to an induced group homomorphism $f_* : H_p(X) \to H_p(Y)$

Functoriality means $I_{X*} = I_{H_p(X)}$ and $(f \circ g)_* = f_* \circ g_*$. A homology theory satisfies other nice properties too, e.g.

- If two maps $f : X \to Y$ and $g : X \to Y$ are homotopic then $f_* = g_*$

The challenge is to set up such a functor ...

Top-level steps in the Jordan-Brouwer proof

- `ISOMORPHIC_HOMOLOGY_GROUPS_EUCLIDEAN_COMPLEMENTS`: If two closed $S, T \subseteq \mathbb{R}^N$ are homeomorphic, their complements have isomorphic homology groups: $H_p(\mathbb{R}^n - S) \cong H_p(\mathbb{R}^n - T)$ for all $p \in \mathbb{Z}$
- `ZEROTH_HOMOLOGY_GROUP`: For any space X, the zeroth homology group $H_0(X)$ is isomorphic to the free abelian group on the set of path components of X. For singular homology with integer coefficients, not necessarily all homology theories
- `ISOMORPHIC_FREE_ABELIAN_GROUPS`: Two free abelian groups are isomorphic iff their generating sets have the same cardinality
- Now if S is homeomorphic to the sphere S^{n-1}, since S^{n-1} is compact so is S and therefore they are both closed so the theorem applies. Their complements are open and $\mathbb{R}^n$ is locally path-connected, so components and path components are the same.

Basic source for the first part

A Simple Proof of the Jordan-Alexander Complement Theorem

Albrecht Dold

The complements of homeomorphic subsets $A, B \subset \mathbb{R}^n$ of Euclidean space need not be homeomorphic, $A \approx B \not\Rightarrow (\mathbb{R}^n - A) \approx (\mathbb{R}^n - B)$. This is well illustrated by classical knot theory, i.e. when A, B are knots in $\mathbb{R}^3$. The complements usually have different fundamental groups in this case, $\pi_1(\mathbb{R}^3 - A) \ncong \pi_1(\mathbb{R}^3 - B)$, and this fundamental group serves to distinguish non-equivalent knots.

On the other hand, it is a classical consequence of **Alexander** duality (cf. **[D]**, VIII, 8.15) that the homology groups of the complements agree if A, B are homeomorphic closed subsets of $\mathbb{R}^n$. Thus,

Theorem. *If $A, B \subset \mathbb{R}^n$ are homeomorphic closed subsets then their complements have isomorphic homology groups, $H(\mathbb{R}^n - A) \cong H(\mathbb{R}^n - B)$,—also in generalised (co-)homology.*

Setting up group theory

We first set up a certain amount of basic group theory (group, homomorphism, isomorphism, quotient group, ...) including the notion of an *exact sequence* of groups and homomorphisms

$$\ldots \xrightarrow{f_{n-1}} G_n \xrightarrow{f_n} G_{n+1} \xrightarrow{f_{n+1}} \ldots$$

meaning that the *image* of each map is the same as the *kernel* of the following map, $\mathrm{im}(f_n) = \ker(f_{n+1})$, where for $f : G \to H$ these are defined by

- $\mathrm{im}(f) = \{f(x) \mid x \in G\}$
- $\ker(f) = \{x \in G \mid f(x) = 1_H\}$

Group theory in HOL Light

Image and kernel (with the source and target groups as extra parameters)

```
|- group_image (G,G') f = IMAGE f (group_carrier G)

|- group_kernel (G,G') f =
     {x | x IN group_carrier G /\ f x = group_id G'}
```

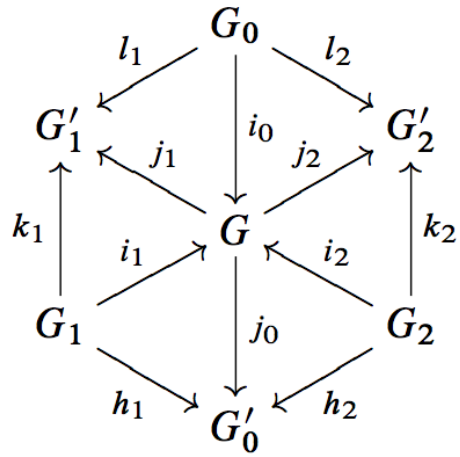
Exactness of a pair of homomorphisms at their 'interface' group:

```
|- group_exactness (G,H,K) (f,g) <=>
       group_homomorphism (G,H) f /\
       group_homomorphism (H,K) g /\
       group_image (G,H) f = group_kernel (H,K) g
```

Commutative diagrams and diagram chasing

There are common patterns involving exact sequences: the *five lemma*, *four lemma*, *snake lemma*, or (Dieck's *Algebraic Topology*):

(11.1.3) Hexagon Lemma. *Given a commutative diagram of abelian groups.*



Suppose that k_1, k_2 are isomorphisms, (i_1, j_2) exact, (i_2, j_1) exact, and $j_0 i_0 = 0$. Then $h_1 k_1^{-1} l_1 = -h_2 k_2^{-1} l_2$.

These are generally easy to prove even formally by "diagram chasing" but *look* ugly and baffling in a purely textual representation.

Free Abelian groups

A basic ingredient in setting singular homology is the notion of a free abelian group over a set S, which is intuitively the set of formal sums

$$a_1 x_1 + \cdots + a_n x_n$$

where $a_i \in \mathbb{Z}$, and $x_i \in S$, with the expected rules for addition, subtraction and multiplication by integers

$$(3x + 4y) - 2(2y - x) = 5x$$

Formally we can regard these as the functions $S \to \mathbb{Z}$ with finite support (i.e. giving 0 for all but finitely many elements of S), with addition etc. pointwise.

Free Abelian groups in HOL Light

In HOL Light we set up a type ‘`:A frag`’ of free Abelian groups over a type `:A`, with associated arithmetic operations `frag_add`, `frag_cmul` etc., and even a handy decision procedure for basic algebraic rearrangements:

```
FRAG_MODULE
 ‘frag_cmul(a + b) c =
  frag_add (frag_cmul a c) (frag_cmul b c)‘;;
```

Linked to the general formalization of group theory like this

```
|- free_abelian_group s =
       group ({c | frag_support c SUBSET s},
              frag_0,
              frag_neg,
              frag_add)
```

Standard simplices

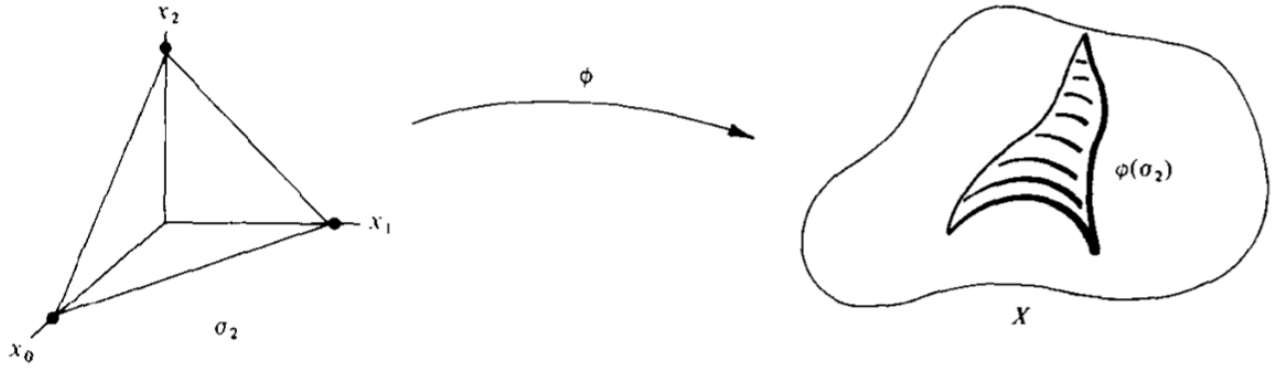
For each $p \in \mathbb{N}$ we define the *standard p-simplex* to be the subset of $\mathbb{R}_{p+1}$ with $\Sigma_0^p x_i = 1$, or in HOL Light:

```
|- standard_simplex p =
   { x:num->real | (!i. &0 <= x i /\ x i <= &1) /\
                   (!i. p < i ==> x i = &0) /\
                   sum (0..p) x = &1}
```

Note that all these live inside the infinite product $\mathbb{R}^{\mathbb{N}}$, not any bounded Euclidean space.

Singular simplices

A *singular* p*-simplex* in a topological space X is simply a continuous function from the standard p-simplex to X:

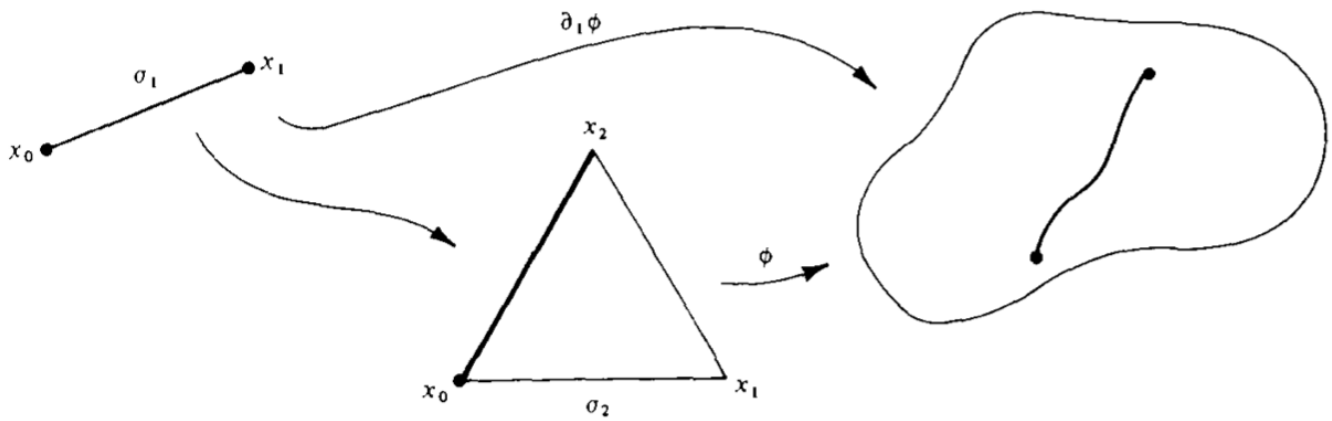


Singular = function need not be injective so image of a singular p-simplex can have topological dimension $< p$. In HOL Light:

```
|- singular_simplex (p,top) f <=>
   continuous_map(subtopology (product_topology (:num)
                                  (\i. euclideanreal))
                              (standard_simplex p),
                  top) f /\
   EXTENSIONAL (standard_simplex p) f
```

Singular face

The singular face map takes a singular p-simplex to a singular $(p-1)$-simplex by composing through the k'th face of the standard simplex




```
|- face_map k x =
     \i. if i < k then x i
         else if i = k then &0 else x(i - 1)

|- singular_face p k f =
       RESTRICTION (standard_simplex (p - 1))
                   (f o face_map k)
```

Singular chains

Now the singular chain group $C_p(X)$ for a topological space X is just the free Abelian group on the singular p-simplices over X:

```
|- singular_chain (p,top) c <=>
       frag_support c SUBSET singular_simplex(p,top)
```

We will often appeal to the fact that functions over the generating set of a free Abelian group automatically extend to the whole group, encoded:

```
|- frag_extend f x =
        iterate frag_add (:A)
          (\a. frag_cmul (dest_frag x a) (f a)))
```

The boundary of a singular chain

We define the boundary of a p-simplex σ to be a sum with alternating signs of the various faces σ_k:

$$\Sigma_{k=0}^{p}(-1)^k\sigma_k$$

This then gives by extension the *boundary homomorphism* $\partial_p : C_p(X) \to C_{p-1}(X)$:

```
|- chain_boundary p c =
       if p = 0 then frag_0 else
       frag_extend
         (\f. iterate frag_add (0..p)
                (\k. frag_cmul (--(&1) pow k)
                       (frag_of(singular_face p k f)))) c
```

Cycles and boundaries

We define the singular cycles $Z_p(X)$ as the kernel of the boundary map ∂_p and the singular boundaries $B_p(X)$ as the image of ∂_{p+1}. More generally we define *relative* forms $Z_p(X, A)$ and $B_p(X, A)$ where we ignore chains in the subtopology for a subset A:

```
|- mod_subset (p,top) c1 c2 <=>
   singular_chain (p,top) (frag_sub c1 c2)

|- singular_relcycle(p,top,s) c <=>
       singular_chain (p,top) c /\
       (chain_boundary p c == frag_0)
          (mod_subset(p-1,subtopology top s))

|- singular_ relboundary(p,top,s) c <=>
     ?d. singular_chain (p + 1,top) d /\
         (chain_boundary (p + 1) d == c)
            (mod_subset (p,subtopology top s))
```

The boundary of a boundary

A fundamental fact (which you can prove by brute force) is that the boundary of a boundary is zero, $\partial_{p-1} \circ \partial_p = 0$:

```
|- singular_chain (p,top) c
   ==> chain_boundary (p - 1) (chain_boundary p c) =
       frag_0
```

This implies every (relative) boundary is a (relative) cycle, $B_p(X, A) \subseteq Z_p(X, A)$:

```
|- singular_relboundary(p,top,s) c
   ==> singular_relcycle(p,top,s) c
```

The (relative) homology relation

At last we define the notion of relative homology: $\sigma, \sigma' \in C_p(X)$ are *homologous relative to* A if their difference is a relative boundary, $\sigma - \sigma' \in B_p(X, A)$:

```
|- homologous_rel (p,top,s) c1 c2 <=>
     singular_relboundary (p,top,s) (frag_sub c1 c2)
```

Intuitively the difference "is the boundary of a region".

Induced map

A continuous map $f : X \to Y$ between topological spaces gives rise to an induced group homomorphism $f_\sharp : C_p(X) \to C_p(Y)$

```
|- simplex_map p g c =
       RESTRICTION (standard_simplex p) (g o c)

|- chain_map p g c =
       frag_extend (frag_of o simplex_map p g) c
```

Homology groups at last!

Define actual groups corresponding to $C_p(X)$ and $Z_p(X, A)$:

```
|- chain_group (p,top) =
   free_abelian_group (singular_simplex (p,top))

|- relcycle_group (p,top,s) =
        subgroup_generated (chain_group (p,top))
           (singular_relcycle (p,top,s))
```

Relative homology group $H_p(X, A)$ is the quotient group $Z_p(X, A)/B_p(X, A)$, and $H_p(X) =_{\mathsf{def}} H_p(X, \emptyset)$.

```
|- relative_homology_group (p,top,s) =
    if p < &0 then singleton_group ARB else
    quotient_group (relcycle_group (num_of_int p,top,s))
      (singular_relboundary (num_of_int p,top,s)))

|- homology_group(p,top) =
     relative_homology_group(p,top,{})
```

Induced map and boundary map

The induced $f_\sharp : C_p(X) \to C_p(Y)$ and boundary $\partial_p : C_p(X) \to C_{p-1}(X)$ homomorphisms appropriately respect the homology relation and can be lifted to the relative homology groups. Informally we write

- $\partial : H_p(X, A) \to H_{p-1}(A)$
- $f_* : H_p(X, A) \to H_p(Y, B)$ when $f : X \to Y$ is a continuous map with $f[A] \subseteq B$

In the actual formalization they are much more heavily parametrized with p and the topological pairs.

Functoriality

The induced map is a functor: identity map inducing identity homomorphism and $(g \circ f)_* = g_* \circ f_*$, or in HOL Light:

```
|- continuous_map (top,top') f /\
   IMAGE f s SUBSET t /\
   continuous_map (top',top'') g /\
   IMAGE g t SUBSET u
   ==> hom_induced p (top,s) (top'',u) (g o f) =
       hom_induced p (top',t) (top'',u) g o
       hom_induced p (top,s) (top',t) f)
```

The heavy parametrization and conditions mean there's often a lot of detail to fill in relative to textbook presentations.

Naturality

The boundary map is a natural transformation, i.e. $\partial \circ f_* = f_* \circ \partial$

```
|- continuous_map (top,top') f /\ IMAGE f s SUBSET t
   ==> hom_boundary p (top',t) o
       hom_induced p (top,s) (top',t) f =
       hom_induced (p - &1) (subtopology top s,{})
         (subtopology top' t,{}) f o
       hom_boundary p (top,s))
```

Again a much more verbose statement informally buried in the assumption that f is a map of topological pairs.

The Eilenberg-Steenrod Axioms

Most applications of homology theory just need a number of basic properties codified by Eilenberg, Steenrod and Milnor. As well as functoriality and naturality:

- If f and g are homotopic then $f_* = g_*$.
- Excision: if $\text{closure}(U) \subseteq \text{interior}(A)$ then the map $\iota_* : H_p(X - U, A - U) \to H_p(X, A)$ induced by inclusion is an isomorphism.
- Dimension: if X is a 1-point space then $H_p(X)$ is the trivial group for $p \neq 0$.
- Exactness of homology sequence

$$\xrightarrow{\partial} H_p(A) \xrightarrow{i_*} H_p(X) \xrightarrow{j_*} H_p(X, A) \xrightarrow{\partial} H_{p-1}(A)$$

- Additivity: added later by Milnor, follows from others for finite sums.

Two of the Eilenberg-Steenrod axioms in HOL Light

The two most difficult of the axioms to prove are homotopy (via a prism construction)

```
|- homotopic_with (\h. IMAGE h s SUBSET t) (top,top') f g
   ==> hom_induced p (top,s) (top',t) f =
       hom_induced p (top,s) (top',t) g
```

and excision (via repeated subdivision of a complex)

```
|- top closure_of u SUBSET top interior_of t /\
   t SUBSET s
   ==> group_isomorphism
         (relative_homology_group
             (p,subtopology top (s DIFF u),t DIFF u),
          relative_homology_group
             (p,subtopology top s,t))
         (hom_induced p
            (subtopology top (s DIFF u),t DIFF u)
            (subtopology top s,t) (\x. x))
```

Applications of homology

HOL Light development is about 5700 lines to get the E-S-M axioms and another 5000 developing consequences of the axioms and making applications:

- The generalized Jordan curve theorem, following Dold (see above)
- Easy development of topological degree from the fact that spheres have (reduced) homology groups isomorphic to $\mathbb{Z}$ so the induced map is just multiplication by an integer, hence other consequences like Brouwer.
- A proof of the Borsuk-Ulam theorem following Dieck's *Algebraic Topology* with a mix of homology and a little bit of homotopy to deform a sphere map.

Lots more interesting stuff to do, e.g. cohomology, homology with other coefficients and the Lefschetz fixed-point theorem, Alexander duality, ...

Thank you!