Notes on Van Kampen's Theorem

Rich Schwartz

September 26, 2022

The purpose of these notes is to shed light on Van Kampen's Theorem. For each of exposition I will mostly just consider the case involving 2 spaces. At the end I will explain the general case briefly. The general case has almost the same proof. My notes will take an indirect approach. I will first explain the construction that lies at the heart of the proof and then I will plug this in to the main argument.

1 Free Products

Let G and H be groups. The *free product* $G * H$ is defined, as a set, to be the the union of the empty word and all words of the form $a_1 a_2 a_3, ...$ with these alternately being nontrivial elements in G and H. The group law is concatenation, followed by all cancellation needed to return to a reduced word. The cancellation process is unique, because we start in the middle, so to speak, and then work our way outward. What happens is that we see aa^{-1} in the middle. We cancel this off and then we have $b_1 b_2$. If this element is nontrivial we replace $b_1 b_2$ by b and stop. Otherwise we cancel off $b_1 b_2$ and repeat.

I can't really improve on Hatcher's proof that $G*H$ is a group, but let me say it somewhat differently, highlighting the main ideas. Inverses are easy: Just reverse the word and replace each element by its inverse. The nontrivial part is showing that the group law is associative.

Let W be the set of reduced words, including the empty word. Let $P(W)$ denote the group of permutations of the *set* W. The main idea is to give an injective map $L : W \to P(W)$ such that $L(ab) = L(a)L(b)$. Since $P(W)$ is a group, it is associative. But then $L((ab)c) = L(a(bc))$. Since L is injective, $(ab)c = a(bc)$. Hence the group law is associative.

Now I give the construction of L. For each element $g \in G \cup H$ the map L_g is defined to be left multiplication: $L_g(a_1 a_2 ...) = g a_1 a_2 ...$ If g, a_1 are in the same group, G or H, we replace ga_1 with the product. If the product is nontrivial we leave it there. Otherwise we remove it. If g, a_1 are not in the same group we leave ga_1 as is. Since $L_{g^{-1}} = L_g^{-1}$ we see that L_g really is a permutation.

If $a_1...a_n$ is a reduced word, we define $L(a_1...a_n) = L_{a_1}...L_{a_n}$, with the product on the right being composition. This defines the map L. Consider $L(ab)$. This differs from $L(a)L(b)$ by cancelling off a finite number of terms of the form $L_g L_{g^{-1}}$, each of which is the trivial permutation. It therefore follows that $L(ab) = L(a)L(b)$. Note that $L(a_1...a_n)$ is nontrivial because this permutation maps the empty string to $a_1...a_n$. So, it is a nontrivial permutation. Hence L is injective.

2 Normal Subgroups of Free Products

Here is the situation we have in Van Kampen's Theorem. We have three groups G, H, K and maps $\phi : K \to G$ and $\psi : K \to H$. These maps need not be injective. We let N be the normal subgroup generated by all words of the form

$$\phi(k)\psi(k)^{-1} \tag{1}$$

for $k \in K$. Note that N is not just the union of these elements. We can also take products as well. Also, conjugation must preserve N, because N is normal. So, N is really the group of finite products of elements conjugate to the kind in Equation 1.

The group we will be interested in is $(G * H)/N$. How is this related to topology? The setup is that A and B are path connected topological spaces such that $A \cap B$ is path connected and furthermore A and B are each open in $A \cup B$. We fix some basepoint $p \in A \cap B$. We have

$$G = \pi_1(A, p), \qquad H = \pi_1(B, p), \qquad K = \pi_1(A \cap B, p).$$

The maps ϕ and ψ are the maps from K to G and from K to H induced by the inclusion from $A \cap B$ respectively to A and to B. Van Kampen's Theorem says that

$$\pi_1(A \cup B, p) = (G * H)/N. \tag{2}$$

3 Decorated Homotopies

Let A, B, p be as above. Suppose that $F : [0,1]^2 \to A \cup B$ is a continous map having the property that $F(0,y) = F(1,y) = p$ for all y. In other words, F maps the left and right edges to P.

Lemma 3.1 *We can divide $[0,1]^2$ into a grid of squares such that F maps each square into either A or into B.*

Proof: This is compactness. Each point of $[0,1]^2$ has a neighborhood which is mapped into either A or B. By compactness we can cover $[0,1]^2$ by finitely many neighborhoods, each of which is mapped into either A or B. Again by compactness, we can divide $[0,1]$ into a sufficiently fine grid so that each square in the grid is contained in one of these special open sets. ♠

We fix this square grid and then we label each square according as to whether F maps the square into A or into B. If F maps a square into $A \cap B$ we just pick randomly. Thus, we have labeled all the squares in the grid by the letters A and B. For the concreteness I will always draw the grid with 9 squares, as in Figure 1, and I will number the squares as indicated.

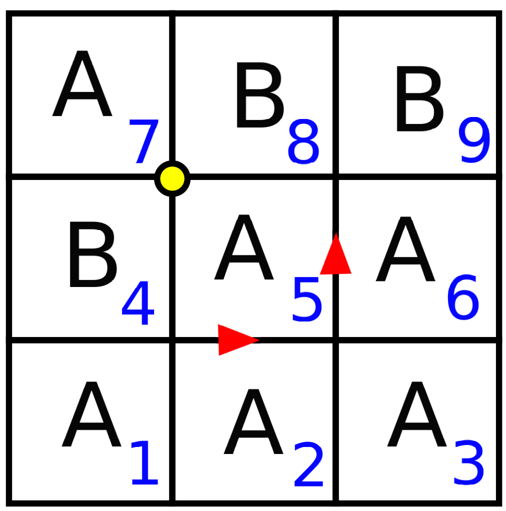

Figure 1: The grid with its face labels.

Each vertex has type AA or BB or AB according as it is surrounded by all A squares, all B squares or both kinds of squares. Likewise each edge has type AA or BB or AB according as it includes in only A-squares, only B-squares, or both.

Lemma 3.2 *Without loss of generality we can assume that F maps each grid vertex to p, the basepoint.*

Proof: Let v be a vertex of the grid. We choose small disks $\Delta_1 \subset \Delta_2$ centered at v and a homeomorphism $h : \Delta_2 - \Delta_1 \to \Delta_1 - p$. We then replace $F|_{\Delta_2 - \Delta_1}$ by $F \circ h$. Now f maps all of $\partial \Delta_1$ to $F(v)$. We now join $F(v)$ to p by a path α and then redefine $F|_{\Delta_1}$ so that $f(\Delta_1) = \alpha$ and $F(p) = v$. The new F sort of stretches Δ_1 along α.

We choose α wisely. The vertex v has type AA or BB or AB according as it is surrounded by all A squares, all B squares or both kinds of squares. If v has type AA the path stays in A. If v has type BB the path stays in B. if v has type AB the path stays in AB. We can do this because A and B and $A \cap B$ are all path connected. This means that the modified F respects the grid structure. ♠

We orient the edges in the picture as indicated. They always point right or up. Let e be an edge and let v_1, v_2 be the endpoints of e, chosen so that e points from v_1 to v_2. Since F maps the endpoints of e to p we can interpret

$$\beta(e) = [F|_e] \tag{3}$$

as an element of the relevant fundamental groups. The edge e has type AA or AB or BB according to the grid squares that are adjacent to it.

- If e has type AA we interpret $\beta(e)$ as an element of $\pi_1(A, p)$.
- If e has type BB we interpret $\beta(e)$ as an element of $\pi_1(B, p)$.
- If e has type AB we interpret $\beta(e)$ *both* as an element of $\pi_1(A, p)$. and as an element of $\pi_1(B, p)$. That is, we associate two group elements to e in this case. Notice that these two elements correspond to the same element of $\pi_1(A \cap B, e)$. They are just defined by the same path. Call such elements *equivalent*. These are the kinds of elements that arise in Equation 1 when we have the setup for Van Kampen's Theorem.

We label each edge by one or two fundamental group elements, according to the scheme above. Each edge gets one or two labels according to its type.

All this structure is what we mean by a *decorated homotopy*.

4 Edge Paths

Suppose now we have a path γ which starts on the left or right edge of $[0,1]^2$ follows finitely many edges, then returns to the left of right edge. To γ we associate a certain finite list of elements in the free product $\pi_1(A,p)*\pi_1(B,p)$. We simply choose one of our edge labels for each edge and write down these labels (or their inverses) in the order we encounter them. If we go forwards along the edge we use this edge label. If we go backwards along the edge we use its inverse. We then do the cancellation to get a reduced word. All the edges along the left and right sides are labeled by the trivial element, so it doesn't matter how much we move up and down along the left and right sides of $[0,1]^2$.

If our path encounters k edges of type AB then we can associate any of 2^k elements in the free product.

Lemma 4.1 *Any two free product elements associated to the same path give the same element in $(\pi_1(A,p)*\pi_1(B,p))/N$ where N is the normal subgroup generated by elements of the form gh^{-1} where g and h are equivalent labels.*

Proof: It suffices to show this for two elements which just differ by one label. Thus we want to compare $U = XgY$ and $V = XhY$ where X, Y are concatenations of labels we don't care about. We have

$$UV^{-1} = XgY(XhY)^{-1} = XgYY^{-1}h^{-1}X^{-1} = X(gh^{-1})X^{-1}.$$

So UV^{-1} is conjugate to gh^{-1} and therefore in N. ♠

Now we define 2 kinds of paths:

1. A *lasso path* is one which starts at $(0,0)$, follows a path to a small square, goes around it, then follows the same path back.

2. The *separator path* σ_k is the one that separates squares $1,...,k$ from squares $k+1,...,n^2$ in the grid. There are n^2 of these, where n^2 is the number of grid squares.

Figure 2 illustrates these kinds of paths. The red path on the right is the separator path σ_4. The paths σ_0 and σ_{n^2} are the paths which go along the top and along the bottom. Here $n = 3$ in our picture.

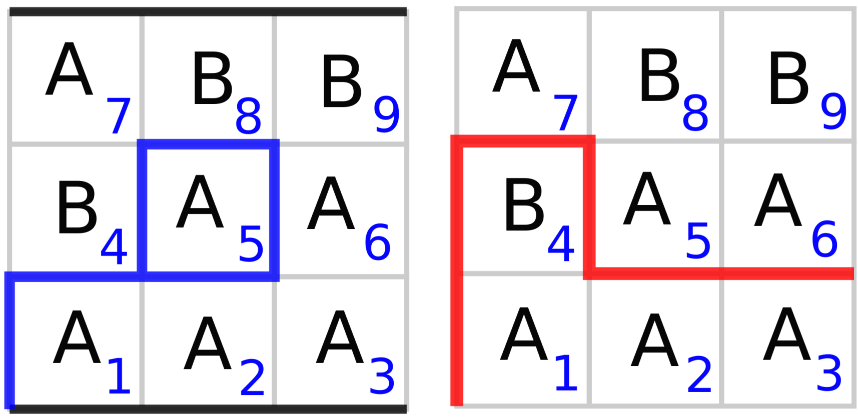

Figure 2: lasso, and separator paths

Lemma 4.2 *Let g be an element in the free product associated to a lasso path. Then $g \in N$, the subgroup from Lemma 4.1.*

Proof: Let's call the lasso loop $\gamma\lambda\gamma^{-1}$ where λ is the loop that goes around the little square. Without loss of generality assume that this square is labeled A. There are potentially many free product labels associated to our lasso, but let us use the one which uses an A-label as much as possible. By Lemma 4.1 our result is true for one of these words if and only if it is true for any other one. So, we might as well prove our result for the maximally-A word.

Since N is a normal subgroup it suffices to prove that $\lambda \in N$. But the word associated to λ is just the product of 4 elements of $\pi_1(A, p)$ corresponding to the boundary of the square. Since F maps the entire square into A, the restriction $F|_A$ induces a homotopy from the word associated to λ and the trivial word. Hence our maximally-A word is in fact the trivial word. The trivial word certainly lies in N. ♠

We suppose that our grid is an $n \times n$ grid. Let $m = n^2$.

Lemma 4.3 *Let g_0 and g_m be the free product words associated to σ_0 and σ_m respectively. Then $g_0 g_m^{-1}$ belongs to the normal subgroup N from Lemma 4.1.*

Proof: We have

$$g_0 g_m^{-1} = (g_0 g_1^{-1})(g_1 g_2^{-1})...(g_{m-1} g_m^{-1}).$$

Here g_j is one of the free product words associated to the separator σ_j. Note that $g_{j-1} g_j^{-1}$ is just a free product word associated to a lasso that goes around square j. Hence this word belongs to N. Since $g_0 g_m^{-1}$ is the product of words that lie in N, it also lies in N. ♠

5 Proof of Van Kampen's Theorem

We have done most of the hard work. Now I'll state the result and its proof. The hypotheses on A and B are as above. Let $i_A^{\#} : \pi_1(A \cap B, p) \to \pi_1(A, p)$ be the map on fundamental groups induced by inclusion. Likewise define $i_b^{\#}$.

Theorem 5.1 (Van Kampen) $\pi_1(A \cup B) = (\pi_1(A, p) * \pi_1(B, p))/N$. *Here N is the normal subgroup generated by elements of the form gh^{-1} where $g = i_A^{\#}(k)$ and $h = i_B^{\#}(k)$.*

Proof: We construct a map from $\pi_1(A, p) * \pi_1(B, p)$ to $\pi_1(A \cup B, p)$. Given some free product element $a_1 b_2 a_3 ...$ we represent each a_i by a loop in A based at p and each b_i by a loop in B based at p. We then concatenate these loops and take the homotopy equivalence class. The same construction works for words starting with b. This map respects concatenation so it is a homomorphism. Call it Ψ.

We just need to check 3 things:

- Ψ is surjective.
- The kernal of Ψ contains N.
- The kernal of Ψ is contained in N.

These three things together show that Ψ induces an isomorphism from the group $(\pi_1(A, p) * \pi_1(B, p))/N$ to the group $\pi_1(A \cup B, p)$.

To see that Ψ is surjective we suppose we are given some loop $f : [0, 1] \to A \cup B$ which maps the endpoints to p. By compactness, we can partition $[0, 1]$ into finitely many intervals such that f_0 maps each interval into either

A or B. This gives us a 1-dimensional version of the grid considered above. We join the vertices of the grid to p as above. This gives us a labeling of the edges of the grid by elements in $\pi_1(A, p)$ and $\pi_1(B, p)$ as above. The product of these words is homotopic to f because the auxilliary paths and their inverses cancel each other out in pairs.

To see that the kernel of Ψ contains N it suffices to show that the generators of N lie in the kernel. That is, we have to show that $\Psi(g) = \Psi(h)$ whenever $g = i_A^{\#}(k)$ and $h = i_B^{\#}(k)$. But both $\Psi(g)$ and $\Psi(h)$ are represented by the same loop, the one that represents k in $A \cap B$.

Now we come to the interesting step. To see that the kernel of Ψ is contained in N suppose that $\Psi(g)$ is trivial. This means that there is a map $F : [0, 1]^2 \to A \cup B$ giving a homotopy between the loop representing g and the constant loop. We can create the decorated homotopy as in the previous sections. Once we do this, one of the free product words associated to the bottom path σ_0 is exactly g. The free product word associated to the top path is the identity. By Lemma 4.3, we have $g \in N$. ♠

6 The General Case

The general case involves the spaces $A_1, ..., A_n$. The hypotheses are that

1. Each A_i is open in the union.
2. All double and triple intersections are path connected.

In this case, Van Kampen's Theorem says that

$$\pi_1(A_1 \cup ... \cup A_n, p) = (\pi_1(A_1, p) * ... * \pi_1(A_n, p))/N \qquad (4)$$

Where N is the normal subgroup generated by all the same elements as above, with respect to every possible pair A_i, A_j.

The proof is the essentially the same. First, we perturb the grid so that at most 3 squares meet at a vertex. Now when we associate paths to the vertices we can keep them inside the relevant triple intersections. Second, we label each square by some A_i rather than an A or a B. Now we just do the same thing and we get the result.