Orienting Transverse Fiber Products

Lyle Ramshaw and Julien Basch

HP Laboratories
HPL- 2009-144

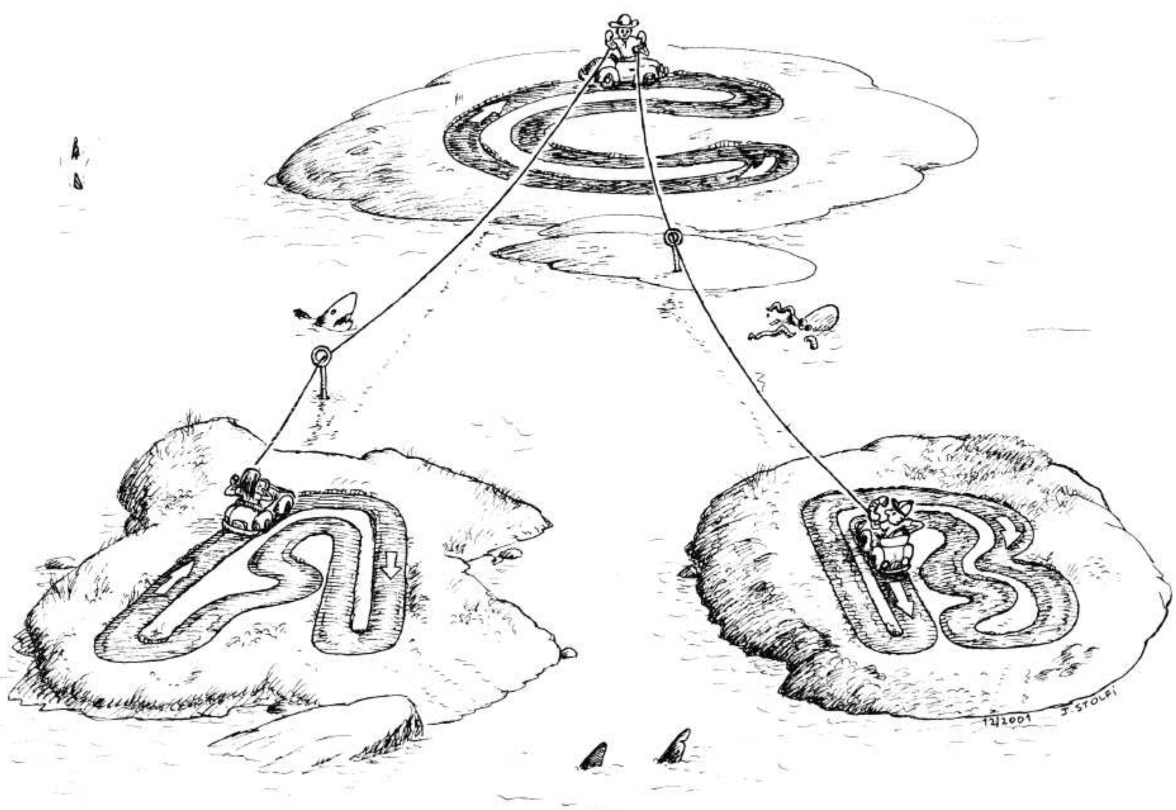
Abstract:
Direct products (a.k.a. Cartesian products) are familiar: Given linear spaces A and B of dimensions a and b, their direct product $A \times B$ is the linear space of dimension $a+b$ consisting of all ordered pairs $(\boldsymbol{a}, \boldsymbol{b})$, for $\boldsymbol{a}$ in A and $\boldsymbol{b}$ in B; and direct products of smooth manifolds are an analogous story. Fiber products (a.k.a. pullbacks) are a less familiar generalization. Given three linear spaces A, B, and S of dimensions a, b, and s and given linear maps $f: A \to S$ and $g: B \to S$, their fiber product $A \times_S B$ is that subspace of the direct product $A \times B$ on which the maps f and g agree, that is, the set of ordered pairs $(\boldsymbol{a}, \boldsymbol{b})$ with $f(\boldsymbol{a}) = g(\boldsymbol{b})$ in S. When the image spaces $f(A)$ and $g(B)$ together span all of S, the maps f and g are said to be *transverse*, and the dimension of the fiber product is then $a+b-s$. Fiber products make sense also for manifolds: Given smooth manifolds A, B, and S of dimensions a, b, and s and given smooth maps $f: A \to S$ and $g: B \to S$ that are transverse in the appropriate sense, it is a standard result that the fiber product $A \times_S B$ is itself a smooth manifold of dimension $a+b-s$.

But what if the input manifolds A, B, and S are oriented? Is there then some natural rule for orienting the fiber-product manifold $A \times_S B$? Such a rule is needed for an application of fiber products in computer-aided geometric design and robotics --- in particular, for computing the boundary of a Minkowski sum from the boundaries of its summands. We show that there is a unique rule for orienting transverse fiber products that satisfies the *Axiom of Mixed Associativity*: $(A \times_S B) \times_T C = A \times_S (B \times_T C)$.

External Posting Date: June 21, 2009 [Fulltext]
Internal Posting Date: June 21, 2009 [Fulltext]

Approved for External Publication

© Copyright 2009 Hewlett-Packard Development Company, L.P.



Orienting Transverse Fiber Products

Extended Abstract

There is a mathematical operation called a “fiber product”, where the word “fiber” here comes from “fiber bundle” — nothing to do with fiber optics, textiles, or constipation. Given sets A, B, and S and given maps $f\colon A \to S$ and $g\colon B \to S$, as in the diagram

$$\begin{tikzcd} A \arrow[dr, "f"'] & & B \arrow[dl, "g"] \\ & S & \end{tikzcd} \qquad (*)$$

the *fiber product* of A and B over S, written $A \times_S B$, is that subset of the Cartesian product $A \times B$ on which the maps f and g agree:

$$A \times_S B := \{(\mathbf{a}, \mathbf{b}) \mid \mathbf{a} \in A, \mathbf{b} \in B, \text{ and } f(\mathbf{a}) = g(\mathbf{b})\}.$$

Note that the fiber product depends upon the maps f and g; when we want to make that dependence explicit, we shall include the maps in the formula for the fiber product, using the nonstandard notation $A[f] \times_S [g]B$.

Fiber products arise in computing the boundary of a Minkowski sum from the boundaries of its summands. Let $\mathcal{A}$ and $\mathcal{B}$ be two regions in the plane (or in 3-space). Their *Minkowski sum* is the region $\mathcal{A} \oplus \mathcal{B} := \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}\}$, where $\mathbf{a} + \mathbf{b}$ here denotes the sum of the points $\mathbf{a}$ and $\mathbf{b}$ as vectors. Each point on the boundary of the Minkowski sum $\mathcal{A} \oplus \mathcal{B}$ is the vector sum of a pair of points, one on the boundary of $\mathcal{A}$ and the other on the boundary of $\mathcal{B}$, where the tangent lines (or tangent planes) are parallel. Finding all such pairs of parallel tangents turns out to be an example of a fiber product: The sets A and B are the boundary curves (or surfaces) of the regions $\mathcal{A}$ and $\mathcal{B}$, the set S is the unit circle (or unit sphere), and the maps f and g are the Gauss maps, the maps that take each boundary point to the outward-pointing unit normal vector at that point.

Where do Minkowski sums arise? Given a region in the plane (or in 3-space), taking its Minkowski sum with a disk (or a ball) corresponds to *offsetting* the boundary curve (or surface) of that region — that is, to moving the boundary a fixed distance orthogonal to itself. Such offsetting arises frequently in computer-aided geometric design (CAGD). Minkowski sums arise also in font design, where one summand models the shape of a brush or pen while the other models the trajectory along which that brush is translated. And Minkowski sums arise in robotics, when computing the configuration space for an object under translation.

In all of these situations, it is common practice to represent the regions involved by specifying their boundaries. So carrying out a Minkowski sum requires computing the boundary of that sum, which involves a fiber product.

Fortunately, there is a standard theory of fiber products that almost suffices. If A, B, and S are smooth d-manifolds and if we rule out various degeneracies by requiring that the smooth maps f and g be *transverse*, then the fiber product $A[f] \times_S [g]B$ is itself a smooth d-manifold.

That standard theory doesn't deal with orientation, however. People doing CAGD or robotics typically orient the boundaries of their regions, to help them distinguish inside from outside. Thus, given the oriented boundaries of regions $\mathcal{A}$ and $\mathcal{B}$, they need the boundary of the Minkowski sum $\mathcal{A} \oplus \mathcal{B}$ also to be oriented. To meet that need, we show how to orient a fiber-product manifold $A[f] \times_S [g]B$, given orientations on the manifolds A, B, and S as input (and always assuming that the maps f and g are transverse). In particular, given points $\mathbf{a}$ in A and $\mathbf{b}$ in B with $f(\mathbf{a}) = g(\mathbf{b})$, we orient the fiber product $A[f] \times_S [g]B$ at the point $(\mathbf{a}, \mathbf{b})$ by working with the Jacobian matrices of f at $\mathbf{a}$ and of g at $\mathbf{b}$.

In the fiber products $A \times_S B$ that arise in computing Minkowski sums, the three input manifolds A, B, and S always have the same dimension. We also tackle the harder problem of orienting the manifold $A \times_S B$ when the dimensions of A, B, and S differ, though we don't know of any practical applications where it is important to orient such fiber products. While such a manifold $A \times_S B$ is clearly orientable, it is no longer at all clear which of its two possible orientations is the proper one — that is, which orientation rule for this more general situation satisfies the most compelling collection of identities. We argue that the proper choice is the unique rule that satisfies $(A \times_S B) \times_T C = A \times_S (B \times_T C)$, a mixed flavor of associativity in which the two fiber products involved on each side are taken over different base manifolds. More precisely, given smooth manifolds and smooth maps of the form

$$\begin{tikzcd}[column sep=small] A \arrow[dr, "f"'] & & B \arrow[dl, "g"] \arrow[dr, "h"'] & & C \arrow[dl, "k"] \\ & S & & T & \end{tikzcd} \qquad (**)$$

the *Axiom of Mixed Associativity* requires that

$$\Big(A[f] \times_S [g]B\Big)[h'] \times_T [k]C = A[f] \times_S [g']\Big(B[h] \times_T [k]C\Big),$$

where the auxiliary maps $h'\colon A \times_S B \to T$ and $g'\colon B \times_T C \to S$ are defined by $h'(\mathbf{a}, \mathbf{b}) := h(\mathbf{b})$ and $g'(\mathbf{b}, \mathbf{c}) := g(\mathbf{b})$. This axiom is so powerful that, together with a few other, more obvious axioms, it determines a unique orientation rule for transverse fiber products.

Acknowledgements

The bulk of this research was done in the 1990s, while I, Lyle Ramshaw, worked at the Systems Research Center of the Digital Equipment Corporation (DEC-SRC) and Julien Basch was a graduate student at Stanford University. DEC was later acquired, first by Compaq, and then by HP.

For an embarrassing span of years, I couldn't bring myself to publish this report, since my investigation of the Axiom of Mixed Associativity suggested some intriguing problems in linear algebra that I was unable to resolve to my satisfaction. I recently learned that other mathematicians have made huge strides in this area by developing the theory of *quiver representations*, to which this report now appeals, when analyzing the structure of zigzags in Section 10.4.1.

I thank my coauthor, Julien Basch, for many key ideas, including the concept of a mixed fiber product, the Axiom of Mixed Associativity, and the clever name "whisker". My thanks go also to Marcos Aguilera, Gunnar Carlsson, Jonathan Derryberry, Leo Guibas, John MacCormick, Jim Saxe, and Jorge Stolfi for lots of help — and to Jorge for the cartoon frontispiece.

Contents

List of Figures

Chapter 1

Alice and Bob

Fiber products are not well known. One way to introduce them is through the story of Alice and Bob, the sweethearts with the neutrino radios.

1.1 Using neutrino radios

Alice has just rowed her boat east, along the Equator, to the western beach of Example Island, whose equatorial cross section is shown in Figure 1.1. Bob has just rowed west to the island's eastern beach. Since they are eager to rendezvous, they shoulder their packs and prepare to walk inland.

Their top priority, though, is to stay in constant contact. They are equipped with the latest in high-tech radios, rumored to operate by transmitting a beam of neutrinos. These radios allow them to stay in constant contact, even when separated by kilometers of solid rock, as long as they remain at the same altitude. (They must also remain over the Equator; they can't wander off north or south.)

They are in contact at the start, since they are both at sea level. By carefully controlling their rates of climb, they manage to stay in contact until the state shown in Figure 1.2. At this point, Alice has to stand still while Bob walks west across his plateau. Then Alice retreats toward the western beach for a while, so that Bob can advance into the valley that next confronts him. Once Bob passes the bottom of his valley, Alice can resume her forward progress, leading to the state shown in Figure 1.3. Bob is almost at his goal — the peak of Mount Tryst, where they hope

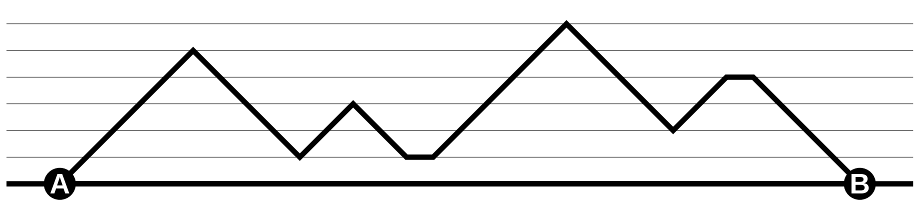


Figure 1.1: Example Island with Alice and Bob at their beaches

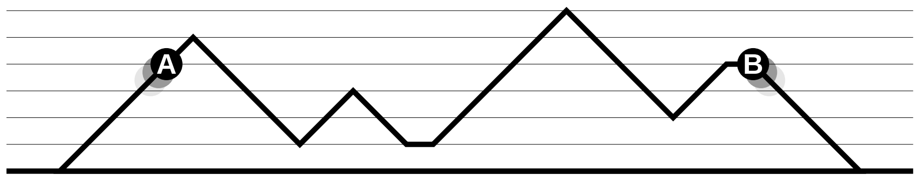


Figure 1.2: Bob reaches the east edge of his plateau

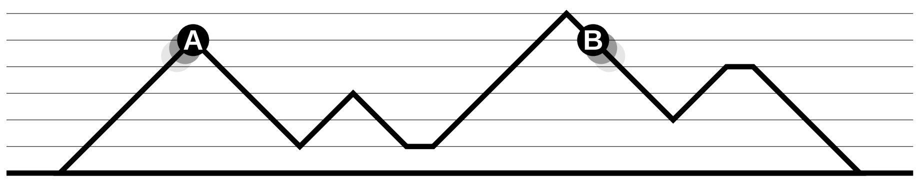


Figure 1.3: Alice at the top of her first hill

to rendezvous soon.

But now it is Bob's turn to retreat, so that Alice can advance into the valley that she now faces. Bob backs up to the bottom of his valley once again. At this point, Alice and Bob have a worried consultation; they decide that their best hope is for both of them to retreat for a while. So Alice climbs westward while Bob climbs eastward, ending as in Figure 1.4. Alice then stands still while Bob retreats across his plateau. Then Bob retreats almost to the eastern beach, allowing Alice to advance to the bottom of her first valley.

Continuing in this fashion, Alice and Bob do eventually manage to rendezvous at the peak of Mount Tryst.

The tilted rectangle $WT \times ET$ in Figure 1.5 represents the state space in which Alice and Bob have their adventure. The lower-left-to-upper-right coordinate of a point in that rectangle represents Alice's west-to-east position in her mountain range — from the western beach W to Mount Tryst T. Thus, starting at any state in the rectangle, we find the corresponding position of Alice by moving diagonally up and left to the boundary of the rectangle and then straight up to Alice's mountain range. The story for Bob is similar: The lower-right-to-upper-left coordinate of a state represents Bob's east-to-west position, from the eastern

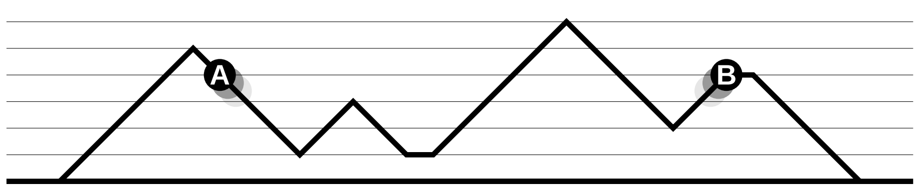


Figure 1.4: Alice and Bob have both been retreating

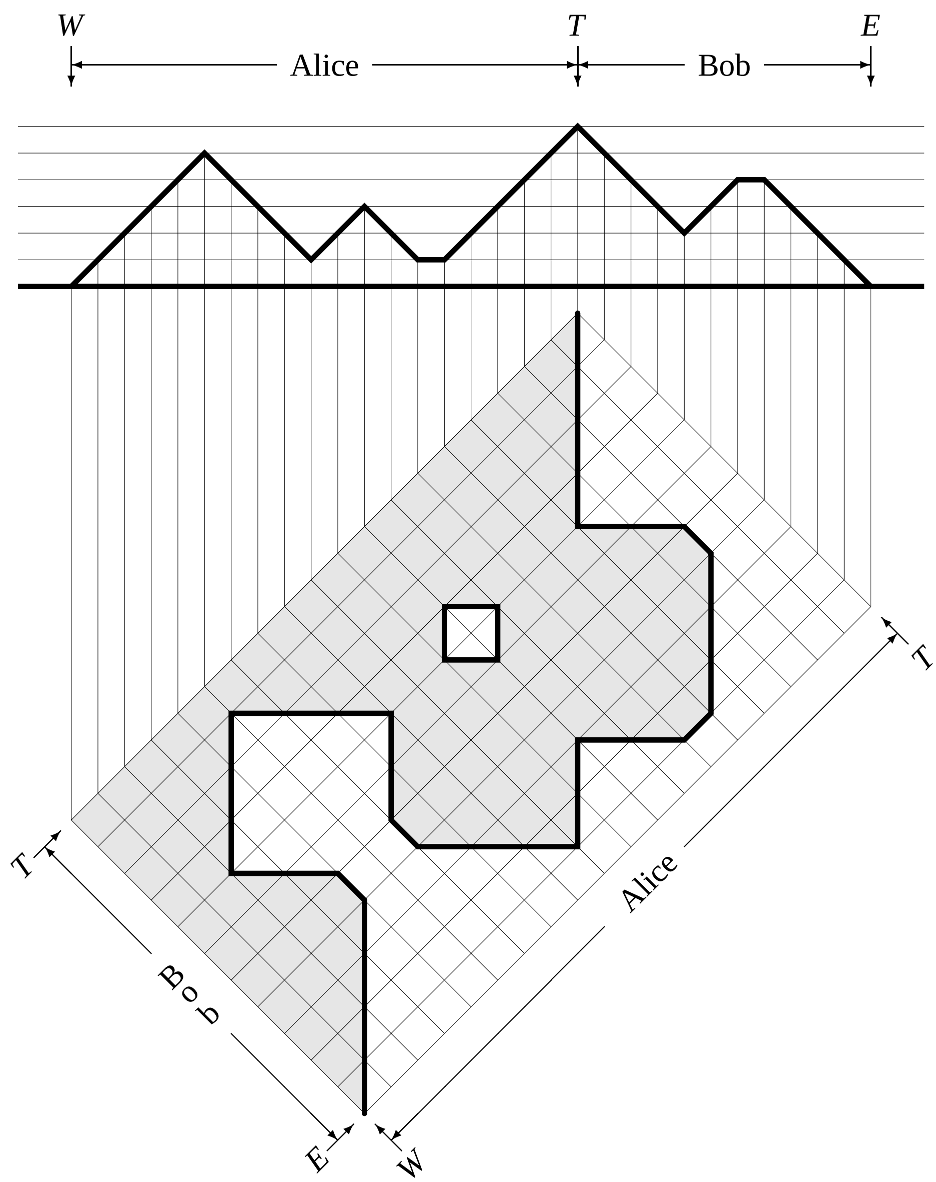


Figure 1.5: The state space of the adventure on Example Island

beach E to Mount Tryst T. Starting at any state in the rectangle, we find the position of Bob by moving diagonally up and right to the boundary of the rectangle and then straight up to Bob's mountain range.

The states in the gray portion of the rectangle $WT \times ET$ are those in which Bob is higher than Alice, while the states in the white portion have Alice higher than Bob. Gray is separated from white by a black path of states in which Alice and Bob are at the same altitude. During their adventure, Alice and Bob follow this black path from the initial state (W, E) to the final state (T, T).

Note that, while the black path goes mostly upward on the page, it winds around enough to include one segment that goes straight down. Alice and Bob had their worried consultation when starting along that segment of the black path.

Note also that there is a little square of white, Alice-higher states sitting in the midst of the gray, Bob-higher region. The boundary of that square is a closed loop of black path, not connected to the long black path that Alice and Bob follow on the way to their rendezvous. This loop arises because Alice's mountain range has a peak that sticks up above the floor of a valley in Bob's range. If Alice is near the top of that peak and Bob is near the bottom of that valley, they can be at the same altitude; and they can climb around a bit while keeping their altitudes equal. But, without losing contact — or tunneling or flying — there is no way for them to get from one component of black path to another.

You may be surprised that the black path can have multiple components in this way. Indeed, multiple components couldn't happen if either of the two mountain ranges were monotonic; but they can happen in the general case, and they do happen in our applications to CAGD and robotics, as we discuss in Section 3.7. In fancier language, a fiber product of connected manifolds may fail to be connected.

1.2 Introducing the fiber product

To model this mathematically, let $A := WT$ denote the set of possible east-west positions of Alice; let $B := ET$ be the same for Bob; and let S denote the set of possible altitudes of either Alice or Bob. We model the topography of Alice's mountain range by an altitude function $f : A \to S$, while we model Bob's range by $g : B \to S$. The rectangle of possible states is the direct (a.k.a. Cartesian) product $A \times B = WT \times ET$. The black path P — which may have multiple components — is the set of pairs $(\mathbf{a}, \mathbf{b})$ in $A \times B$ for which the altitudes are equal:

$$P := \{(\mathbf{a}, \mathbf{b}) \in A \times B \mid f(\mathbf{a}) = g(\mathbf{b})\}.$$

This set P is called the *fiber product* of the maps f and g.

The term "fiber" comes from a mathematical structure called a "fiber bundle", by the way. Given two fiber bundles over the same base space S, the appropriate way to multiply them is to take their fiber product over S, which will be another fiber bundle over S. Fiber products arise also in category theory, where they are often called *pullbacks*.

Fiber products can be defined in many contexts, as we discuss in Section 4.4. We shall typically assume that the spaces A, B, and S are compact smooth manifolds without boundary and that the maps $f: A \to S$ and $g: B \to S$ are smooth. But the situation that Alice and Bob face is somewhat different.

For one thing, Alice's mountain range A is the closed segment WT. This segment is a smooth 1-manifold and is compact, but it does have boundary points: to wit, the endpoints W and T. The story for Bob's range $B = ET$ is similar. Boundary points can cause trouble, as we discuss in Section 1.5.

A second thing that is different about Alice and Bob's situation is that the altitude functions f and g in Figure 1.5 are piecewise affine[1], rather than smooth. But we don't change anything essential if we round off the peaks and valleys slightly, so as to make them smooth. When we do so, the corners of the fiber product P also get rounded off, converting P itself into a smooth 1-manifold. (There is no need to round off Mount Tryst; that is, the derivative of f from the west at T need not agree with the derivative of g from the east at T. Indeed, it is better not to round off Mount Tryst, for reasons that we discuss below.)

We can define an altitude function $h: P \to S$ on the fiber product P by letting $h(\mathbf{a}, \mathbf{b})$, for any state $(\mathbf{a}, \mathbf{b})$ on the black path, denote the common altitude $h(\mathbf{a}, \mathbf{b}) := f(\mathbf{a}) = g(\mathbf{b})$ of Alice and Bob in that state. When taking the fiber product of smooth maps $f: A \to S$ and $g: B \to S$, the resulting map $h: P \to S$ is also smooth. To be precise, it is actually the smooth map h that is the *fiber product* of the smooth maps f and g, written $h = f \times_S g$. That is, the objects that really get multiplied, when we take a fiber product over a smooth manifold S, are smooth maps from smooth manifolds to S. We shall hence refer to the maps f and g as the *factor maps* of the fiber product.

The fiber-product manifold P is the domain of the fiber-product map $h = f \times_S g$. In this monograph, we shall write P using the nonstandard notation $P = A[f] \times_S [g]B$. It will often be clear from the context, however, which smooth maps $f: A \to S$ and $g: B \to S$ are intended. We then abbreviate the manifold P by writing simply $P = A \times_S B$, which is the standard notation. But keep in mind that the manifold $A \times_S B$ depends upon the factor maps f and g, even when their names are elided.

1.3 Transversality

Figure 1.6 illustrates some phenomena that didn't arise in Figure 1.5.

First, Alice's and Bob's mountain ranges have peaks, at U and at V, that are at the same altitude. Suppose that Alice and Bob are in the black state (U, V). When they leave that state, each of them can decide independently whether to descend to the west or to the east. Thus, there are four black paths that lead away from the state (U, V) in Figure 1.6, rather than only two.

[1] Functions like f and g are often called "piecewise linear", rather than "piecewise affine"; but I prefer to reserve the word "linear" for things that are homogeneous, as well as of degree 1.

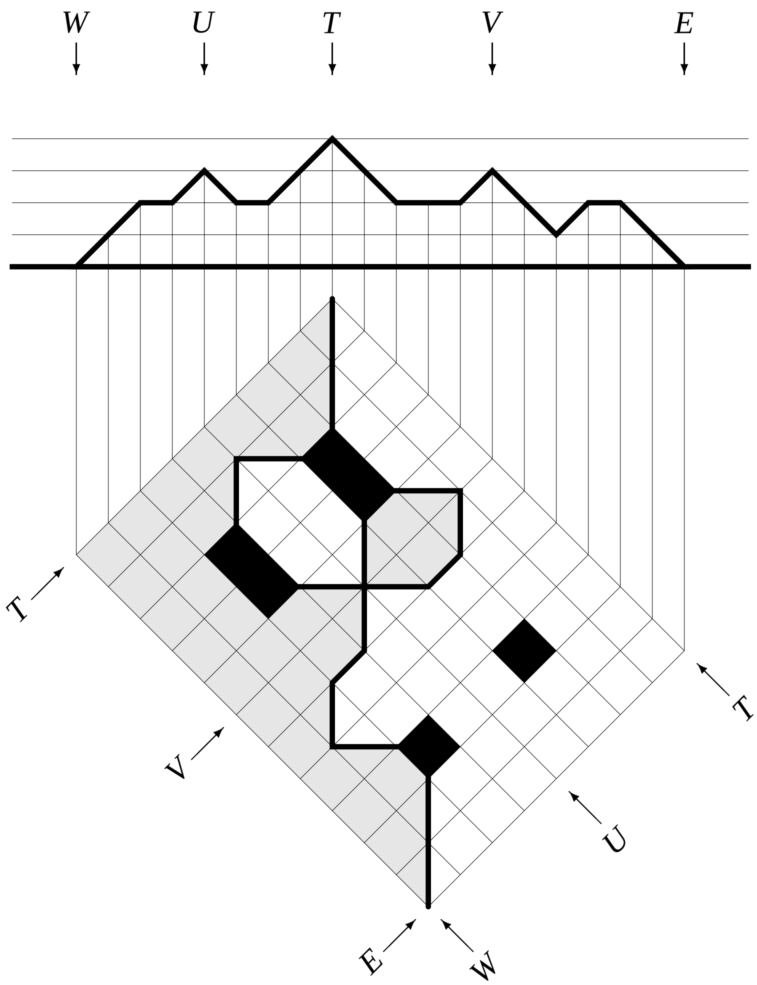


Figure 1.6: An island whose mountain ranges are not transverse

The solid black rectangles in Figure 1.6 are an even more obvious sign of a new phenomenon. They arise because these two mountain ranges have plateaus at the same altitude. Alice and Bob can wander around freely and independently on such a pair of plateaus while staying in constant contact. Thus, the fiber product P — the black region that separates gray from white — includes entire solid rectangles. Such rectangles can have black paths attached either to all four of their corners, to two adjacent corners, to two opposite corners, or to none of their corners, depending upon whether each of the two plateaus involved is a mesa, a bench, or a playa. (A mountain peak is a zero-length mesa, while a valley is a zero-length playa. Thus, the state (U, V) in Figure 1.6, which is a zero-by-zero black rectangle with black paths attached to all four of its corners, results from Alice's and Bob's ranges having equal height, zero-length mesas at U and at V.)

The fiber product in Figure 1.6 is not a 1-dimensional manifold, both because of what happens in the neighborhood of the state (U, V) and because it includes solid rectangles. There is a standard technical condition, called *transversality*, that outlaws the degeneracies in Figure 1.6, hence guaranteeing that the fiber product will be a manifold. In particular, whenever the maps f and g are *transverse*, the fiber product $P = A[f] \times_S [g]B$ is a smooth manifold of dimension $\dim(P) = \dim(A) + \dim(B) - \dim(S)$, as we discuss in Section 4.7.

We also discuss in Section 4.7 what it means for the two maps f and g to be transverse, which is a bit subtle in higher dimensions. Roughly speaking, some combination of moving the point $\mathbf{a}$ around in A and moving $\mathbf{b}$ around in B must cause either $f(\mathbf{a})$ or $g(\mathbf{b})$ to cover all of the dimensions of S. In the case of Alice and Bob, where the manifolds A, B, and S are all 1-dimensional, this boils down to a simple condition: There is only one dimension in S to cover, and Alice covers it by herself unless she is at a flat spot, that is, unless $f'(\mathbf{a}) = 0$. Bob covers it by himself unless he is at a flat spot, with $g'(\mathbf{b}) = 0$. So trouble arises only when Alice and Bob are simultaneously at flat spots. Avoiding trouble means arranging that none of the flat spots on Alice's range are at the same altitude as any of the flat spots on Bob's range. Thus, two factor maps f and g are transverse in the 1-dimensional case just when[2] there are no equal-height flat spots, that is, just when there are no points $(\mathbf{a}, \mathbf{b})$ in the state space $A \times B$ with $f(\mathbf{a}) = g(\mathbf{b})$ and $f'(\mathbf{a}) = g'(\mathbf{b}) = 0$. Recall that we are rounding off all corners (except for Mount Tryst); so the bad state (U, V) in Figure 1.6 does have $f'(U) = g'(V) = 0$. (If we rounded off the peak of Mount Tryst, the resulting flat spot at T would, all by itself, cause transversality to fail at the destination state (T, T). That is why it is better not to round off Mount Tryst — although losing transversality only at a boundary point of the state space might be tolerable.)

1.4 Which way to go?

Let's return to the transverse case: either to the Example Island in Figure 1.5 or to some other island on which Alice's and Bob's mountain ranges have no equal-height flat spots. The fiber product P is then a 1-dimensional manifold. We now consider the problem of orienting P.

To clarify the problem, suppose that Alice and Bob made camp last night. They just woke up this morning. They can talk to each other on their radios, they can look around at their local terrain, and they can tell — say, from where the sun is rising — which way is east. But they can't remember which way they were walking when they stopped to make camp last night. (Such confusion is not unreasonable, since each of them may have walked back and forth past their current location many times already — although there was no single time in the past when they were both in their current locations simultaneously.) Is there some rule by which Alice and Bob can figure out which way to start walking today?

To put that question another way, consider the arrowheads in Figure 1.7. The

[2] I use the phrases "just when" and "precisely when" to mean "if and only if".

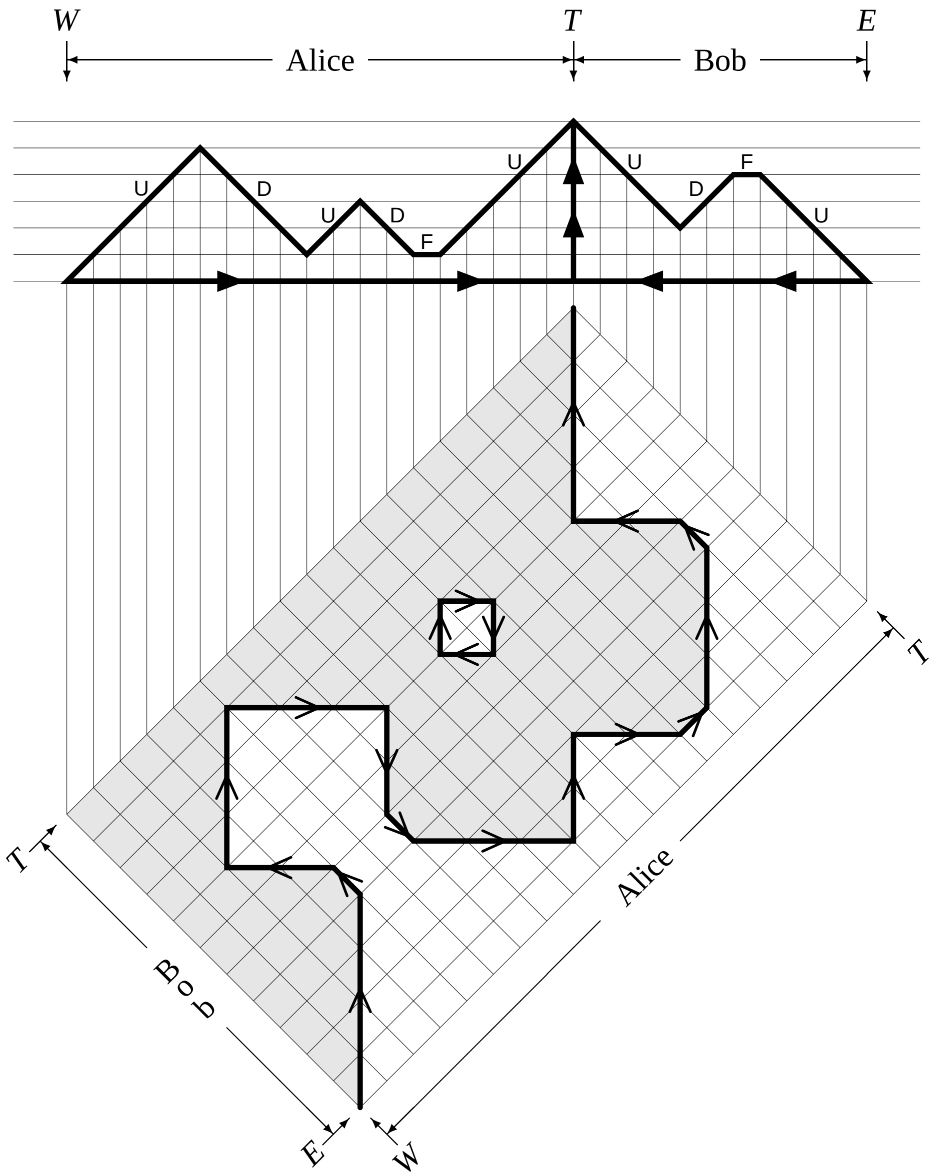


Figure 1.7: The black path for Example Island, correctly oriented

solid arrowheads on $A = WT$, $B = ET$, and S indicate the preferred orientations on those 1-manifolds, while the open arrowheads on the fiber-product 1-manifold P indicate the way that Alice and Bob should walk. Is there some local rule that computes the open arrowheads from the solid ones? A rule is *local* when its decisions are based solely on the slopes where Alice and Bob are currently standing, without exploiting global information about the topography of the island.

Note that there are open arrowheads also on the border of the little white square in the middle of the gray region. Rules that orient fiber products typically orient every component of them, even though Alice and Bob follow only one of those components. Different components could be oriented independently, in principle. But we are studying local orientation rules, rules that are based solely on the slopes. Any local rule must orient the left edge of the little white square in the same way that it orients the vertical segment above the starting point (W, E), since the states along those two segments look locally just the same: In all of them, Alice's terrain slopes up to the right, while Bob's slopes up to the left.

It turns out to be convenient to measure the slopes on Bob's mountain range backwards. In particular, Alice and Bob agree that

advancing means moving toward Mount Tryst, following the horizontal, solid arrowheads in Figure 1.7, while

retreating means moving back toward your own beach.

So Alice advances by moving east, while Bob advances by moving west. Alice and Bob then measure the slope of their local terrain as the rate of change in their altitude that would result from a small advance. So the slope of Alice's terrain is just the slope of the graph of her altitude function f, as in a calculus class; but the slope of Bob's terrain is the negative of the slope of g. Each segment of terrain in Figure 1.7 is marked with U, D, or F according as that terrain is upward-sloping, downward-sloping, or flat — that is, according as Alice or Bob measures the slope to be positive, negative, or zero.

1.4.1 The Greedy-Alice Rule

We are going to discuss two local rules that compute the correct open arrowheads. But before we do that, let's briefly discuss one that doesn't, the *Greedy-Alice Rule*:

> Alice should advance — unless Bob's terrain is flat, in which case Alice must stand still and Bob should advance.

Figure 1.8 shows the open arrowheads generated by the Greedy-Alice Rule. The difficulty is that they are inconsistent; they don't agree with each other about the direction in which the black path should be oriented. (The isolated little square of black path has four open arrowheads on it, and they don't agree either.) Such inconsistencies aren't allowed on an oriented manifold, as we discuss in Section 5.6.

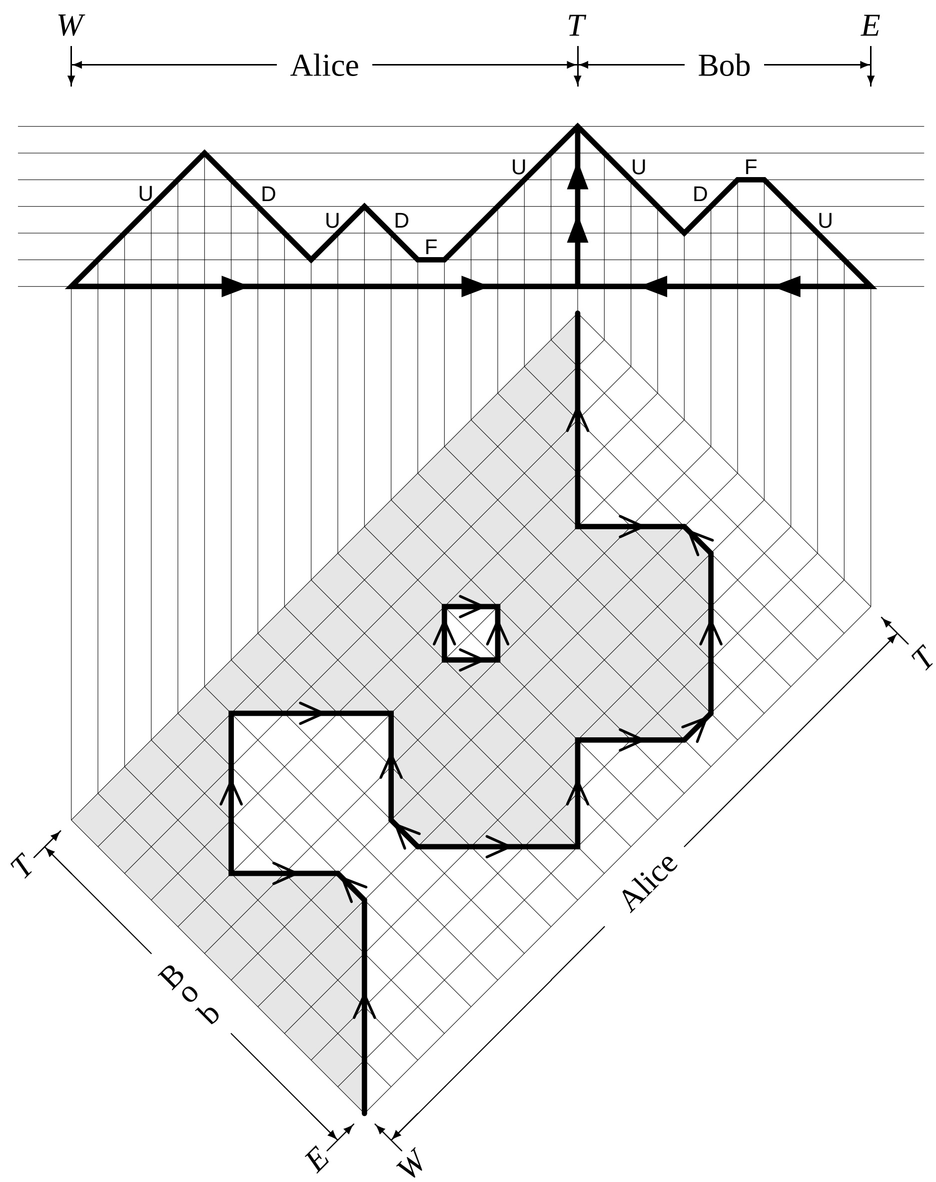


Figure 1.8: The black path oriented inconsistently by the Greedy-Alice Rule

The problem with the Greedy-Alice Rule — the reason that it can generate open arrowheads that are inconsistent — is essentially a lack of continuity. The Greedy-Alice Rule bases its decisions on the slopes of Alice's and Bob's terrains, so it is local. But a tiny change to those slopes can cause the Greedy-Alice Rule to suddenly reverse its recommended open arrowhead. For example, suppose that Alice's slope is $+1$ while Bob's slope is approximately zero. If Bob's slope is slightly positive, then Bob should advance at a sprightly pace, so that his altitude will increase, hence allowing Alice to advance a trifle. On the other hand, if Bob's slope is slightly negative, then he should retreat at a sprightly pace, again so that Alice can advance a trifle. So a tiny change in Bob's slope causes a drastic change to his recommended motion. It is this lack of continuity that leads to the inconsistent open arrowheads in Figure 1.8. We define a notion of continuity for local orientation rules in Section 6.2 that we call *stability*; and we show, in Section 6.3, that any local orientation rule that is stable always generates open arrowheads that are consistent.

Once the open arrowheads are consistent, they are either consistently correct or consistently incorrect, since the black path has only two orientations: the one that we want, which goes from the point (W, E) toward (T, T), and its opposite.

1.4.2 The Partner's-Slope Rule

To develop a rule that is both consistent and correct, let's retrace the first few steps along the black path. At the starting point (W, E) in Figure 1.7, both Alice and Bob are on upward slopes and they should both advance. When Bob reaches his plateau, as was shown back in Figure 1.2, Alice has to stand still. And when Bob moves onto his downward slope, Alice starts to retreat. Thus, Alice's direction of motion seems to be controlled by the slope of Bob's terrain. Symmetry suggests that Bob's motion should be controlled by Alice's slope. That is, both Alice and Bob should employ the following *Partner's-Slope Rule*:

- If your partner's terrain slopes upward, you advance.
- If your partner's terrain slopes downward, you retreat.
- If your partner's terrain is flat, you stand still.

The first step in analyzing this Partner's-Slope Rule is to check that it gives compatible instructions to Alice and to Bob: Whenever Alice is instructed to go uphill, Bob had better be instructed to go uphill also, and the same for going downhill and for staying at the same altitude. To verify compatibility, it suffices to consider the eight possible cases for Alice's and Bob's slopes, as shown in Figure 1.9. (Seven of those eight cases occur in Figure 1.7, the missing case being the one in which Alice's terrain is flat and Bob's slopes downward.) Why are there only eight possible cases, instead of nine? The case that is ruled out is for Alice and Bob to be simultaneously at flat spots. If they ever found themselves in such

Bob \ Alice	U	F	D
U	Alice advances Bob advances (both go up)	Alice advances Bob stands	Alice advances Bob retreats (both go down)
F	Alice stands Bob advances		Alice stands Bob retreats
D	Alice retreats Bob advances (both go down)	Alice retreats Bob stands	Alice retreats Bob retreats (both go up)

Figure 1.9: The eight cases of the Partner's-Slope Rule

a state, the Partner's-Slope Rule would instruct them both to stand still, so they would be stuck. But such states are forbidden by our assumption of transversality.

The second step in the analysis is to check stability. Since our formal definition of stability is still some chapters away, let's here argue informally that the open arrowheads will be consistent. The danger of inconsistency arises when we move from one of the eight cases in Figure 1.9 to an adjacent case. For example, as we move down the leftmost column, Alice first advances, then stands, then retreats. But we move down that column because Bob's slope changes sign from positive to negative. If Bob's slope is only slightly positive, then Alice can advance only quite slowly. Thus, if Bob's slope subsequently changes to become slightly negative, Alice's conversion to a slow retreat constitutes a continuous change in her velocity, not a sudden and discontinuous reversal of orientation. Note that Bob continues to advance at a sprightly pace throughout this process; we don't get the sudden reversal from a sprightly advance to a sprightly retreat that we saw in the Greedy-Alice Rule. The other three sides of the square in Figure 1.9 are similar.

Once we have convinced ourselves that the Partner's-Slope Rule gives open arrowheads that are consistent, it is easy to see that they are consistently correct. We simply check that we get the correct answer in the starting state, when Alice and Bob are at their beaches. They are then both on upward slopes, and we do want them both to advance, as the Partner's-Slope Rule specifies.

Exercise 1-1 Here is a simple rule that tells Alice and Bob, not only the direction in which to walk, but precisely how fast: They each set their own signed velocity to be their partner's signed slope. This *Velocity Variant* of the Partner's-Slope Rule unifies the eight separate cases of Figure 1.9 through the magic of multiplying signed numbers. Unfortunately, this clever idea does not generalize to higher dimensions, for reasons discussed in Exercise 2-8.

In detail, Alice's slope is the rate of change of her altitude $f(\mathbf{a})$ with respect to her position $\mathbf{a}$, which is $df/d\mathbf{a}$. Bob's is $dg/d\mathbf{b}$. Suppose that Alice controls her position $\mathbf{a}$, as a function of time t, so as to make $d\mathbf{a}/dt = dg/d\mathbf{b}$, while Bob makes $d\mathbf{b}/dt = df/d\mathbf{a}$. Show that their rates of climb will then be identical.

Hint: By the Chain Rule, Alice's rate of climb df/dt is then given by

$$\frac{df}{dt} = \frac{df}{d\mathbf{a}}\frac{d\mathbf{a}}{dt} = \frac{df}{d\mathbf{a}}\frac{dg}{d\mathbf{b}}.$$

1.4.3 The Gray-Region Rule

If Alice and Bob are willing to reason about their 2-dimensional state space, that is, about the rectangle $A \times B = WT \times ET$ in Figure 1.7, then they can compute the correct open arrowheads by using the following *Gray-Region Rule*:

> Proceed along the black path in the direction that causes the gray region to lie to your left.

This Gray-Region Rule is simpler than the Partner's-Slope Rule in several ways. For one thing, it unifies the eight separate cases in Figure 1.9: Alice and Bob always leave the gray region to their left, regardless of whether their slopes are upward, downward, or flat. Also, the Gray-Region Rule is obviously stable: As we move along the boundary of the gray region, keeping it always on our left, the open arrowheads that we produce are obviously consistent.

But is the Gray-Region Rule local? That is, can Alice and Bob implement it without having global knowledge about their altitude functions $f: A \to S$ and $g: B \to S$? Yes, they can. Suppose that they are currently at some point $(\mathbf{a}, \mathbf{b})$ on the fiber product, so that $f(\mathbf{a}) = g(\mathbf{b})$. Alice and Bob can measure the slopes of their local terrains. By conferring over their radios, they can combine those slopes to determine, to first order, which points (X, Y) near $(\mathbf{a}, \mathbf{b})$ have $f(X) < g(Y)$ and which have $f(X) > g(Y)$. A first-order sketch of the neighborhood of the point $(\mathbf{a}, \mathbf{b})$ in the direct product $A \times B$ will look like one of the eight sketches in Figure 1.10. In each sketch, the origin represents the point $(\mathbf{a}, \mathbf{b})$; moving to the right means that Alice advances, while moving up means that Bob advances. The gray region is the set of states (X, Y) for which $f(X) < g(Y)$. For example, consider the upper-left sketch, in which both Alice and Bob are on upward slopes. Any combination of Bob advancing and Alice retreating leads to states (X, Y) with $f(X) < f(\mathbf{a}) = g(\mathbf{b}) < g(Y)$; so the entire second quadrant is gray, along with portions of the first and third quadrants. The other sketches are similar. Given any such local sketch, the Gray-Region Rule tells Alice and Bob which way to go; so the Gray-Region Rule is indeed local.

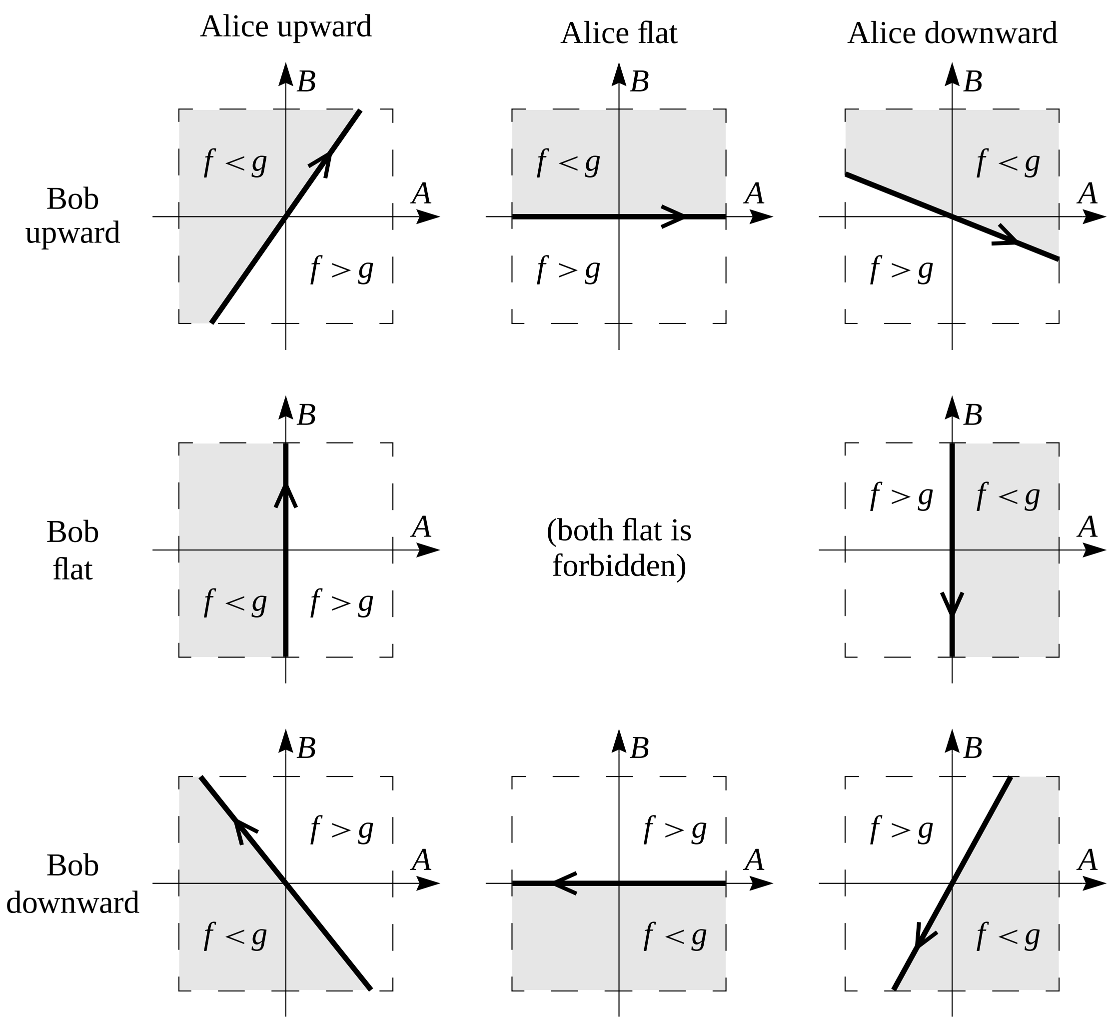


Figure 1.10: The eight cases for the local comparison of $f(X)$ to $g(Y)$

Of course, we must still avoid the non-transverse case, the central case in Figure 1.10, in which both Alice and Bob are at flat spots. It makes sense for Alice and Bob to leave the gray region to their left only if they are on a segment of black path that separates gray from white. It wouldn't make sense, for example, if they were in the middle of the one of the solid black rectangles in Figure 1.6.

The stability of the Gray-Region Rule is geometrically evident in Figure 1.10. As we transition among the eight cases, say cycling clockwise around the figure, the black path rotates continuously — also clockwise, as the figure happens to be drawn. And the arrowheads that the Gray-Region Rule recommends for that black path also rotate continuously, with no sudden reversals. That continuity is the property that we formalize in Section 6.2 as *stability*. Note that the central, non-transverse case has to be outlawed in order to allow this continuity; there are pairs of slopes arbitrarily close to $(0, 0)$ for which the recommended arrowhead points in any specified direction.

By comparing Figure 1.10 to Figure 1.9, we find that the Partner's-Slope Rule and the Gray-Region Rule give the same answers in all eight cases. Of course, it was clear from the start that they must, since they both give the correct answers.

1.4.4 The Intrinsic Gray-Region Rule

Warning: The Gray-Region Rule as stated above would fail if Alice and Bob chose to draw their state space in a different, but equally valid, way. Figure 1.11 shows Example Island again, but with the state space $A \times B$ drawn above the island, rather than below it. The state space in Figure 1.11 is the mirror image of the one in Figure 1.7, reflected in the horizon. Because of this reflection, Alice and Bob should leave the gray region to their right in Figure 1.11, rather than to their left.

We can rephrase the Gray-Region Rule more intrinsically, so that it gives the correct answers in both figures. Let α be a vector (technically speaking, in the tangent space $T_{(\mathbf{a},\mathbf{b})}(A \times B)$ to the manifold $A \times B$ at the point $(\mathbf{a}, \mathbf{b})$) that points in the Alice-advance direction; and let β point in the Bob-advance direction. Rotating from α to β defines an orientation on the 2-dimensional state space $A \times B$. That orientation is counterclockwise in Figure 1.7, but clockwise in Figure 1.11. Let φ be a vector that points along the fiber product $P = A \times_S B$ in the direction in which Alice and Bob should walk. And let δ be a vector that points somewhere into the gray region; that is, the vector δ should point toward some nearby state (X, Y) for which $f(X)$ and $g(Y)$ are distinct points in the manifold S and for which the motion in S from $f(X)$ to $g(Y)$ agrees with the given orientation on S. The *Intrinsic Gray-Region Rule* thens tells us:

> The vector φ orients the fiber product $A \times_S B$ correctly just when the orientation on the direct product $A \times B$ given by rotating from α to β is the same as that given by rotating from φ to δ.

Exercise 1-2 The next day, Alice and Bob explore Example Island again, but

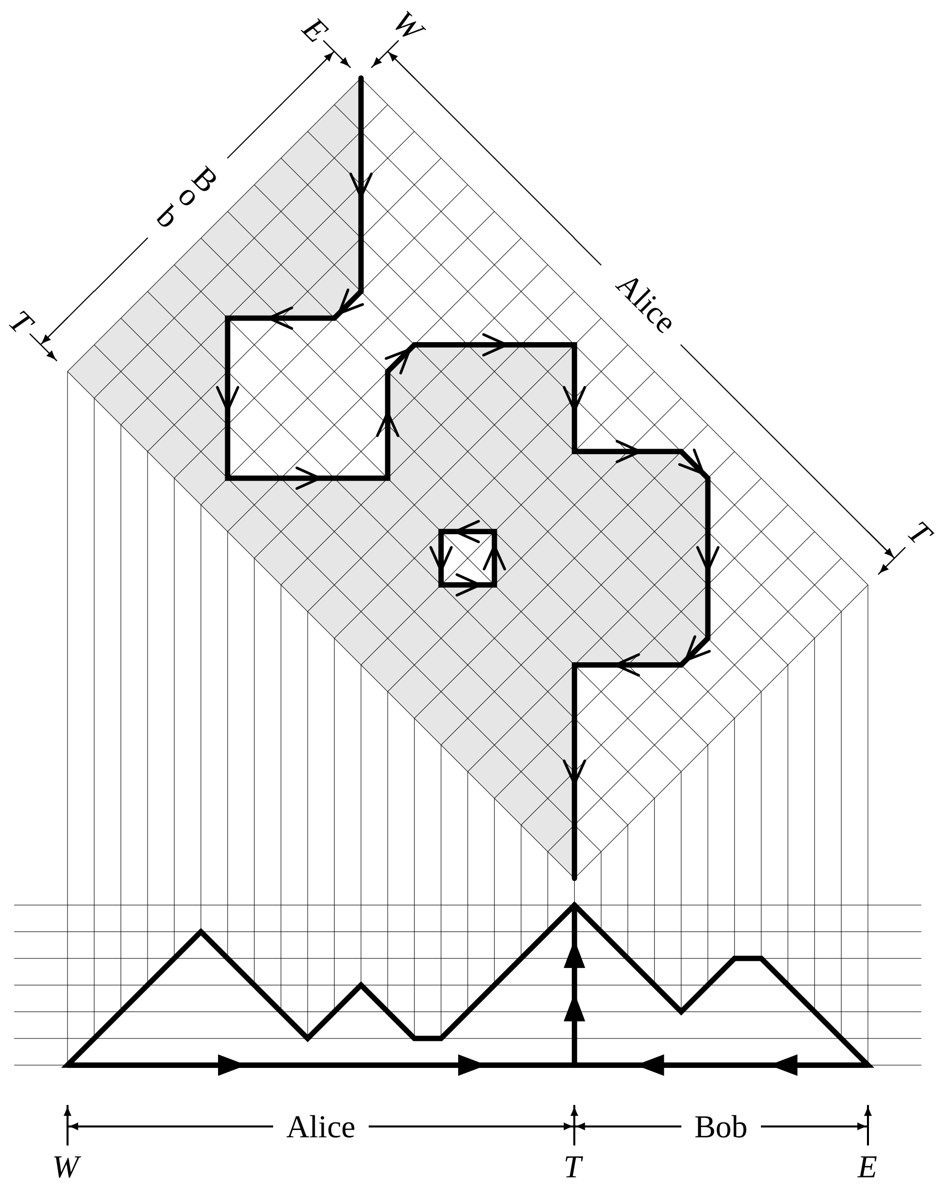


Figure 1.11: Example Island with its state space drawn above it

with Alice walking in from the east this time, while Bob walks in from the west. Explain why the Intrinsic Gray-Region Rule gives the correct answer in this case.

Answer: One way to draw the state space for the second day's adventure is to take Figure 1.7, to swap the labels "Alice" and "Bob" wherever they appear, and to swap the colors of the gray and white regions. Using primes to indicate the vectors on the second day, we then have $\alpha' = \beta$, $\beta' = \alpha$, and $\delta' = -\delta$; so we get $\varphi' = \varphi$. Thus, the fiber product $B \times_S A$ gets the same orientation as $A \times_S B$.

1.5 Farewell to Alice and Bob

We discussed Alice and Bob to develop our intuitions about fiber products. We are about to move on, to consider fiber products of manifolds of higher dimension. But we owe Alice and Bob some closing remarks. In particular, can they always rendezvous, regardless of the topography of their island?

1.5.1 Valleys below sea level

Even in the transverse case, they can fail to rendezvous if one of their mountain ranges, say Alice's, has a valley that dips below sea level. In that case, Alice and Bob will end their adventure with Alice stuck at sea level on the west side of that valley, unable to advance because Bob has retreated all the way to the eastern beach — and he has no scuba gear. In the analog of Figure 1.5, the black path that starts at the initial state (W, E) will leave the rectangle $WT \times ET$ at some point interior to its lower-right side. To avoid this problem, we henceforth assume that every point interior to the island lies strictly above sea level. (We could allow interior points precisely at sea level, but it simplifies things to forbid them.)

Note that this scuba problem arises because Alice and Bob are walking on 1-manifolds with boundary; in particular, the eastern beach E is a boundary point of Bob's manifold. Fortunately, in our applications of fiber products to CAGD and robotics, the manifolds involved are manifolds without boundary.

1.5.2 Guaranteed success

By combining all of our explicit and implicit assumptions, we have arrived at a flavor of island on which Alice and Bob are guaranteed to succeed. Suppose that all interior points of the island are above sea level; that the peak of Mount Tryst is the single highest point on the island; that the altitude function $f\colon A \to S$ of the portion of the island west of Mount Tryst is C^1, that is, continuously differentiable; similarly, for the altitude function $g\colon B \to S$ of the portion east of Mount Tryst; and that the two altitude functions f and g are transverse — that is, there are no states $(\mathbf{a}, \mathbf{b})$ in $A \times B$ with $f(\mathbf{a}) = g(\mathbf{b})$ and $f'(\mathbf{a}) = g'(\mathbf{b}) = 0$. Then, Alice and Bob are guaranteed to succeed, and here is why.

Consider the boundary of the state-space rectangle $A \times B$, as in Figure 1.5. The vertices (W, E) and (T, T) are black. The states $(\mathbf{a}, E)$ for $\mathbf{a} > W$ are all white, since all interior points of Alice's range are above sea level. The states $(T, \mathbf{b})$ for $\mathbf{b} < T$ are also white, since Mount Tryst is the single, highest peak. So the two right-facing sides of the state-space rectangle are entirely white. By a similar argument, the two left-facing sides are entirely gray. Thus, the only two points where a black path can reach the boundary of the state-space rectangle are at its bottom and top corners, the points (W, E) and (T, T).

It follows from transversality that the white and gray regions are separated by a 1-dimensional manifold of class C^1, possibly with multiple components. A component of that black manifold is either a closed loop or an arc diffeomorphic to a closed interval; in the latter case, both endpoints of the arc must lie on the boundary of the rectangle. Thus, there is precisely one component in the black manifold of the latter type: a C^1 black arc that leads Alice and Bob from the initial state (W, E) to the final state (T, T).

By the way, we are exploiting the assumption that Mount Tryst is the single highest peak in a way that might not be obvious. Because of that assumption, we can restrict Alice to walk only west of Mount Tryst, while Bob walks only east of it — and those restrictions are crucial for achieving transversality. It would be more straightforward to let both Alice and Bob walk anywhere on the entire island $WE := WT \cup TE$, hence using the full square $WE \times WE$ as our state space, rather than just the rectangular subset $WT \times TE$. But then any flat spot on the entire island would be enough to destroy transversality, since Alice and Bob could stand, arm in arm, at that flat spot.

1.5.3 The horrors of nontransversality

The assumptions in Section 1.5.2 arguably make things too easy for Alice and Bob. They are walking along an unbranched path in the state space, so it is always obvious what they should do next. To make their adventure more challenging, it is tempting to allow nontransversality. Unfortunately, that opens the door to various monsters from real analysis, one of which is shown in Figure 1.12. Note first that the altitude functions f and g in that figure are not transverse; the left endpoint of the bumpy region on Bob's range is a flat spot that is at the same altitude as the plateau on Alice's range. While the functions f and g are not transverse, they are C^1 — and might even be C^∞, if the bumpy region in Bob's mountain range behaves like the function $x \mapsto e^{-1/x^2} \sin(1/x)$ behaves on the interval $[0 \,..\, 1]$. Alice has to be east of her plateau in order for Bob to advance over any one of the hills in his bumpy region, while Alice has to be west of her plateau in order for Bob to advance through any of the dales. Since Bob faces an infinite sequence of alternating hills and dales, Alice has to cross her plateau infinitely often before they can rendezvous. (Figure 1.12 is a smoothed version of a bad example in Whittaker [16]. Huneke [7] later showed that this example is, in a sense, the only bad thing that nontransversality allows.)

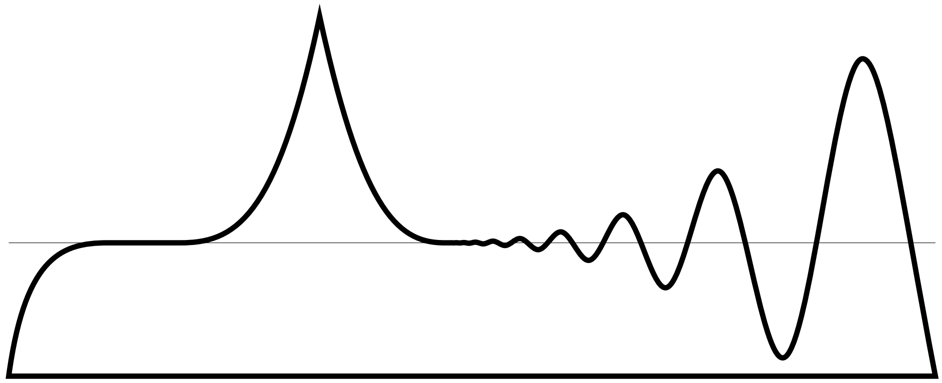


Figure 1.12: An island on which Bob's range is monstrously bumpy

To avoid this type of monstrous behavior without requiring transversality, we would have to require that our altitude functions be even nicer than C^∞, in some way. For example, Whittaker [16] analyzed altitude functions whose graphs can be partitioned into a finite number of strictly monotonic segments. A different way to achieve much the same effect is to require that the altitude functions be piecewise real analytic. These cases are more challenging for Alice and Bob, since their "black path" can have the bad features that appear in Figure 1.6. But Whittaker showed that Alice and Bob can still rendezvous.

Unfortunately, it is not at all clear how to generalize beyond the transverse case in a fiber product $A[f] \times_S [g]B$ of higher dimension. To avoid requiring that the maps f and g be transverse, we must require that they be nicer than C^∞, in some sense. But in what sense? Pierre Schapira [15] has done exciting work in this area, exploiting the notion of a *subanalytic stratification*. Very briefly, a *semianalytic set* is a set that can be described by finitely many equalities and inequalities among real-analytic functions, just as a *semialgebraic set* is one that can be described by finitely many equalities and inequalities among polynomials [2]. While semialgebraic sets behave nicely under projections (the Tarski-Seidenberg Theorem), semianalytic sets do not. A *subanalytic set* is a more subtle and more permissive notion than a semianalytic set, and this extra permissiveness restores nice behavior under projections [2]. Perhaps monstrous behavior can be ruled out without insisting on transversality by requiring that the graphs of the factor maps f and g be subanalytic sets.

We leave such questions as topics for future research. In this monograph, we take the coward's way out and simply require that the factor maps f and g in our fiber products $A[f] \times_S [g]B$ be transverse. This guarantees that our fiber products will be smooth manifolds, which we take it as our challenge to orient.

Exercise 1-3 Sketch the state space of the monstrous island in Figure 1.12.

Chapter 2

Fiber products in higher dimensions

> Convention: We shall often be analyzing a smooth map from one smooth manifold to another. In the neighborhood of any given point, such a map is approximated to first order by a linear map between the tangent spaces. From now on, let's use boldface letters to name the smooth manifolds and smooth maps, while using italic letters to name the approximating linear spaces[1] and linear maps.

It is a standard result that the transverse fiber products of smooth manifolds are themselves smooth manifolds. We are looking for a local rule that orients the output of such a fiber product, given orientations on its inputs. If we were willing to restrict ourselves to those fiber products $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$ in which the three manifolds $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{S}$ are all 1-dimensional, we could stop now. In that 1-*dimensional case*, both the Partner's-Slope Rule and the Gray-Region Rule are local rules that compute the correct open arrowheads, so we are already done twice over. But we want to tackle several generalizations: the

equidimensional case, where $\dim(\mathbf{A}) = \dim(\mathbf{B}) = \dim(\mathbf{S})$, but that common dimension can be any nonnegative integer; and the

any-dimensional case, where the three dimensions are unconstrained — though the problems of most interest have $\dim(\mathbf{A}) + \dim(\mathbf{B}) \geq \dim(\mathbf{S})$, since the only way to achieve transversality when $\dim(\mathbf{A}) + \dim(\mathbf{B}) < \dim(\mathbf{S})$ is vacuously, resulting in an empty fiber product.

The equidimensional case is the one that arises when computing Minkowski sums in CAGD and robotics. It is intriguing to consider the any-dimensional case also, even though we have no practical applications of that case in mind. So what happens when we try to generalize our two rules from the 1-dimensional case to the equidimensional and any-dimensional cases?

[1] Linear spaces are often called "vector spaces". I prefer the name "linear space" because the elements of such a space are often covectors, tensors, matrices, or functions, rather than vectors.

2.1 Local orientation rules

The Partner's-Slope Rule and the Gray-Region Rule are local rules; given any point $(\mathbf{a}, \mathbf{b})$ on the smooth manifold $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$, they tell us how that fiber product should be oriented in the neighborhood of $(\mathbf{a}, \mathbf{b})$ based solely on the local behavior of the maps $\mathbf{f}$ and $\mathbf{g}$, that is, based solely on Alice's slope $\mathbf{f}'(\mathbf{a})$ and Bob's slope $\mathbf{g}'(\mathbf{b})$. We want our rules for higher dimensions to be local as well. But, once the dimensions of the smooth manifolds $\mathbf{A}$ and $\mathbf{S}$ exceed 1, the local behavior of a smooth map $\mathbf{f}\colon \mathbf{A} \to \mathbf{S}$ near a point $\mathbf{a}$ in $\mathbf{A}$ can no longer be described by giving a single real number, the slope $\mathbf{f}'(\mathbf{a})$. Instead, we need an entire matrix, called the *Jacobian*. We next review some standard facts about tangent spaces, differentials, and Jacobian matrices.

2.1.1 Tangent spaces and differentials

A smooth manifold can be approximated to first order near any of its points by a linear space called the *tangent space*. We denote the tangent space to the manifold $\mathbf{A}$ at the point $\mathbf{a}$ as $T_{\mathbf{a}}\mathbf{A}$.

A smooth map between smooth manifolds can be approximated to first order by a linear map between the appropriate tangent spaces. Let $\mathbf{f}\colon \mathbf{A} \to \mathbf{S}$ be a smooth map between smooth manifolds, and suppose that the point $\mathbf{a}$ in $\mathbf{A}$ is carried, by $\mathbf{f}$, to the point $\mathbf{s} := \mathbf{f}(\mathbf{a})$ in $\mathbf{S}$. The behavior of $\mathbf{f}$ near the point $\mathbf{a}$ is then approximated, to first order, by a linear map from $T_{\mathbf{a}}\mathbf{A}$ to $T_{\mathbf{s}}\mathbf{S}$ called the *differential* of $\mathbf{f}$ at $\mathbf{a}$. We denote the differential as $T_{\mathbf{a}}\mathbf{f}\colon T_{\mathbf{a}}\mathbf{A} \to T_{\mathbf{s}}\mathbf{S}$; the notations $(d\mathbf{f})_{\mathbf{a}}$ and $\mathbf{f}'(\mathbf{a})$ are also used. The *Jacobian* of $\mathbf{f}$ at $\mathbf{a}$ is the matrix of the differential $T_{\mathbf{a}}\mathbf{f}$, expressed in terms of chosen bases for the two tangent spaces that it relates.

We shall often use a and s to denote the dimensions of the smooth manifolds $\mathbf{A}$ and $\mathbf{S}$. So the tangent space $T_{\mathbf{a}}\mathbf{A}$ is an a-dimensional linear space, while $T_{\mathbf{s}}\mathbf{S}$ is s-dimensional. The Jacobian of $\mathbf{f}$ at $\mathbf{a}$ is thus an s-by-a matrix. In the special case $a = s = 1$, this matrix reduces to a single number: Alice's slope $\mathbf{f}'(\mathbf{a})$.

Abusing notation even further, we shall often use A to denote the tangent space $A := T_{\mathbf{a}}\mathbf{A}$ itself, when the context makes clear which manifold $\mathbf{A}$ and point $\mathbf{a}$ in that manifold are intended. We similarly abbreviate $S := T_{\mathbf{s}}\mathbf{S}$. And we use f to denote the differential $f := T_{\mathbf{a}}\mathbf{f}$ of the smooth map $\mathbf{f}$ at $\mathbf{a}$. Thus, the smooth map $\mathbf{f}\colon \mathbf{A} \to \mathbf{S}$ is approximated to first order near the point $\mathbf{a}$ by the differential $T_{\mathbf{a}}\mathbf{f}\colon T_{\mathbf{a}}\mathbf{A} \to T_{\mathbf{s}}\mathbf{S}$, which we abbreviate as the linear map $f\colon A \to S$. The Jacobian of $\mathbf{f}$ at $\mathbf{a}$ is the matrix of this linear map f, with respect to chosen bases for A and S; and we denote that matrix as $[f]$.

2.1.2 Orientation

We also need to review briefly what it means to orient a manifold; for more details, see Chapter 5.

Orienting a smooth manifold makes sense in any dimension. For 1-manifolds, the issue is forward versus backward; for 2-manifolds, it is clockwise versus counterclockwise; for 3-manifolds, it is right-handed versus left-handed. We orient a manifold by orienting all of its tangent spaces in some locally consistent manner. (Local consistency isn't always possible; some manifolds, such as the Möbius strip, are not orientable.) And we orient a linear space, such as a tangent space, by assigning a sign to each ordered basis in one of the two possible globally consistent ways. Consistency means that two ordered bases for the same space must be assigned the same sign just when the square matrix that expresses the first basis in terms of the second (or vice versa, it doesn't matter) has positive determinant. Otherwise, the determinant of the change-of-basis matrix will be negative, and the two bases must be assigned opposite signs.

Let $\mathbf{f}\colon \mathbf{A} \to \mathbf{S}$ be a smooth map between smooth manifolds, and suppose that $\mathbf{f}$ is a local diffeomorphism between some neighborhood of the point $\mathbf{a}$ in $\mathbf{A}$ and some neighborhood of the point $\mathbf{s} := \mathbf{f}(\mathbf{a})$ in $\mathbf{S}$. This requires that the two manifolds $\mathbf{A}$ and $\mathbf{S}$ be of the same dimension $a = s$. If the manifolds $\mathbf{A}$ and $\mathbf{S}$ are also oriented, the local diffeomorphism provided by $\mathbf{f}$ near $\mathbf{a}$ must either preserve or reverse orientation. Let's set

$$\operatorname{sense}(\mathbf{f}, \mathbf{a}) := +1 \qquad \text{or} \qquad \operatorname{sense}(\mathbf{f}, \mathbf{a}) := -1$$

according as $\mathbf{f}$ preserves or reverses orientation at $\mathbf{a}$. It is a standard result that the smooth map $\mathbf{f}$ is a local diffeomorphism at $\mathbf{a}$ just when its differential there, the linear map $T_{\mathbf{a}}\mathbf{f}$, is invertible. Recall that we are abbreviating that linear map $T_{\mathbf{a}}\mathbf{f}\colon T_{\mathbf{a}}\mathbf{A} \to T_{\mathbf{s}}\mathbf{S}$ as the map $f\colon A \to S$. We set $\operatorname{sgn}(f) := \pm 1$ according as the linear bijection $f\colon A \to S$ preserves or reverses orientation, so that we have $\operatorname{sgn}(f) = \operatorname{sense}(\mathbf{f}, \mathbf{a})$.

We can concretely test the sign of a linear map by considering its matrix. Let $[f]$ be the matrix of the differential f, which is the Jacobian of $\mathbf{f}$ at $\mathbf{a}$. And let's suppose that, in writing out the matrix $[f]$, we have chosen bases for the tangent spaces A and S that are positively oriented. We can then determine whether $\mathbf{f}$ preserves or reverses orientation at $\mathbf{a}$ by testing the sign of the determinant of its Jacobian: $\operatorname{sense}(\mathbf{f}, \mathbf{a}) = \operatorname{sgn}(f) = \operatorname{sgn}(\det([f]))$.

2.1.3 The maps of a smooth-manifold fiber product

Figure 2.1 sets up some notation for the various maps associated with the smooth-manifold fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$. The two bottom sides of the diamond are the factor maps $\mathbf{f}\colon \mathbf{A} \to \mathbf{S}$ and $\mathbf{g}\colon \mathbf{B} \to \mathbf{S}$. In the Alice-and-Bob case, they measured altitude. The two top sides of the diamond are projections, the maps $\mathbf{u}\colon \mathbf{A} \times_{\mathbf{S}} \mathbf{B} \to \mathbf{A}$ given by $\mathbf{u}(\mathbf{a}, \mathbf{b}) := \mathbf{a}$ and $\mathbf{v}\colon \mathbf{A} \times_{\mathbf{S}} \mathbf{B} \to \mathbf{B}$ given by $\mathbf{v}(\mathbf{a}, \mathbf{b}) := \mathbf{b}$. Note that the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ is the largest subset of the direct product $\mathbf{A} \times \mathbf{B}$ on which $\mathbf{f}\circ\mathbf{u}(\mathbf{a}, \mathbf{b}) = \mathbf{g}\circ\mathbf{v}(\mathbf{a}, \mathbf{b})$, that is, on which $\mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b})$. So the diamond in Figure 2.1 commutes. The diagonal of the diamond, the map $\mathbf{f} \times_{\mathbf{S}} \mathbf{g}\colon \mathbf{A} \times_{\mathbf{S}} \mathbf{B} \to \mathbf{S}$, is a symmetric name for the common composition $\mathbf{f} \circ \mathbf{u} = \mathbf{g} \circ \mathbf{v}$.

$$
\begin{tikzcd}
 & \mathbf{A} \times_{\mathbf{S}} \mathbf{B} \arrow[dl, "\mathbf{u}"'] \arrow[dr, "\mathbf{v}"] \arrow[dd, "\mathbf{f} \times_{\mathbf{S}} \mathbf{g}"] & \\
\mathbf{A} \arrow[dr, "\mathbf{f}"'] & & \mathbf{B} \arrow[dl, "\mathbf{g}"] \\
 & \mathbf{S} &
\end{tikzcd}
$$

Figure 2.1: Five maps associated with the fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$

We abbreviate the dimensions of the manifolds **A**, **B**, and **S** as a, b, and s. So it takes a coordinates to describe the position of Alice, a point **a** in **A**, and b coordinates to describe a point **b** in **B**. As Alice and Bob vary their positions, they want to remain at the same altitude, where it now takes s coordinates to describe an altitude. That is, they constrain their joint motion so as to preserve the altitude equality $\mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b})$, which in turn encodes s scalar equalities. The transverse case is the case in which those s scalar equalities are independent, so it costs a full s degrees of freedom to preserve them all. The transverse fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$ therefore has dimension $a + b - s$.

Exercise 2-1 Consider the fiber-product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$ in Figure 2.1. If the factor map **g** is injective, show that the projection **u** is also injective. Similarly, if **g** is surjective, show that **u** is also surjective. The maps **f** and **v** are related similarly. (By the way, this has nothing to do with smooth manifolds; it holds equally well when **A**, **B**, and **S** are arbitrary sets and **f** and **g** are arbitrary set maps.)

Answer: For injectivity, consider two points $(\mathbf{a}, \mathbf{b}_1)$ and $(\mathbf{a}, \mathbf{b}_2)$ in the fiber product that project to the same point **a** under **u**. We have $\mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b}_1) = \mathbf{g}(\mathbf{b}_2)$. If **g** is injective, it follows that $\mathbf{b}_1 = \mathbf{b}_2$, and hence **u** is injective.

For surjectivity, let **a** be any point in **A**. If **g** is surjective, there must exist **b** in **B** with $\mathbf{g}(\mathbf{b}) = \mathbf{f}(\mathbf{a})$. The point $(\mathbf{a}, \mathbf{b})$ then lies in the fiber product and projects to **a**, so the map **u** is surjective.

2.1.4 The maps of a linear-space fiber product

Consider the behavior of the fiber product shown in Figure 2.1 near some point $(\mathbf{a}, \mathbf{b})$, where $\mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b})$; and let **s** denote that common point $\mathbf{s} := \mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b})$ in **S**. The manifolds **A**, **B**, and **S** are approximated, near the points **a**, **b**, and **s**, by the tangent spaces $T_{\mathbf{a}}\mathbf{A}$, $T_{\mathbf{b}}\mathbf{B}$, and $T_{\mathbf{s}}\mathbf{S}$, which we abbreviate as A, B, and S. Figure 2.2 shows the linear maps relating these spaces — the spaces and maps that approximate Figure 2.1 to first order near $(\mathbf{a}, \mathbf{b})$.

The factor map $\mathbf{f}\colon \mathbf{A} \to \mathbf{B}$ is approximated to first order, near the point **a**, by its differential, the linear map $T_{\mathbf{a}}\mathbf{f}\colon T_{\mathbf{a}}\mathbf{A} \to T_{\mathbf{s}}\mathbf{S}$, which we abbreviate as $f\colon A \to S$. In a similar way, we abbreviate the differential $T_{\mathbf{b}}\mathbf{g}\colon T_{\mathbf{b}}\mathbf{B} \to T_{\mathbf{s}}\mathbf{S}$ simply as $g\colon B \to S$. Note that the map $f = T_{\mathbf{a}}\mathbf{f}$ is invertible just when **f** is a

$$\begin{tikzcd}
& A \times_S B \arrow[dl, "u"'] \arrow[dr, "v"] \arrow[dd, "f \times_S g"] & \\
A \arrow[dr, "f"'] & & B \arrow[dl, "g"] \\
& S &
\end{tikzcd}$$

Figure 2.2: The behavior of the fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$ near $(\mathbf{a}, \mathbf{b})$

local diffeomorphism at **a**, mapping some neighborhood of **a** diffeomorphically onto some neighborhood of **s**.

When the smooth maps **f** and **g** are transverse, it is a standard theorem that the fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$ is itself a manifold. And the tangent space to that fiber product, at the point $(\mathbf{a}, \mathbf{b})$, is simply the fiber product of the tangent spaces; that is, we have the canonical isomorphism

$$T_{(\mathbf{a},\mathbf{b})}(\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}) \cong T_{\mathbf{a}}\mathbf{A}[T_{\mathbf{a}}\mathbf{f}] \times_{T_{\mathbf{s}}\mathbf{S}} [T_{\mathbf{b}}\mathbf{g}]T_{\mathbf{b}}\mathbf{B}.$$

Leaving the factor maps implicit, we have $T_{(\mathbf{a},\mathbf{b})}(\mathbf{A} \times_{\mathbf{S}} \mathbf{B}) \cong T_{\mathbf{a}}\mathbf{A} \times_{T_{\mathbf{s}}\mathbf{S}} T_{\mathbf{b}}\mathbf{B}$; or, using our italic-letter convention, we have simply $T_{(\mathbf{a},\mathbf{b})}(\mathbf{A} \times_{\mathbf{S}} \mathbf{B}) \cong A \times_S B$.

Finally, consider the projection map $\mathbf{u}\colon \mathbf{A} \times_{\mathbf{S}} \mathbf{B} \to \mathbf{A}$ from Figure 2.1, which is given by $\mathbf{u}(\mathbf{a}', \mathbf{b}') := \mathbf{a}'$. Near the point $(\mathbf{a}, \mathbf{b})$, this map **u** is approximated to first order by its differential, which is a linear map $T_{(\mathbf{a},\mathbf{b})}\mathbf{u}\colon T_{(\mathbf{a},\mathbf{b})}(\mathbf{A} \times_{\mathbf{S}} \mathbf{B}) \to T_{\mathbf{a}}\mathbf{A}$ that we shall abbreviate as $u\colon A \times_S B \to A$. This linear map u is simply the projection of linear spaces, the map given by $u(\alpha, \beta) := \alpha$, for all vectors α in A and β in B with $f(\alpha) = g(\beta)$. In a similar way, the differential of the projection $\mathbf{v}\colon \mathbf{A} \times_{\mathbf{S}} \mathbf{B} \to \mathbf{B}$ at $(\mathbf{a}, \mathbf{b})$ is the projection $v\colon A \times_S B \to B$ given by $v(\alpha, \beta) := \beta$.

Thus, if we start with the fiber product of smooth manifolds and smooth maps shown in Figure 2.1 and we consider that fiber product to first order near the point $(\mathbf{a}, \mathbf{b})$, what we end up with is the fiber product of linear spaces and linear maps shown in Figure 2.2. When we say that an orientation rule for fiber products of smooth manifolds is *local*, we require that it decide on an orientation for the manifold $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ near the point $(\mathbf{a}, \mathbf{b})$ solely by working with the linear spaces and linear maps in Figure 2.2. Thus, any local rule for orienting smooth-manifold fiber products uses some rule for linear-space fiber products as a subroutine; the smooth-manifold rule works by applying that subroutine, independently, to the tangent spaces and differentials at each point.

2.2 Generalizing the Partner's-Slope Rule

With those concepts added to our toolkit, let's try to generalize the Partner's-Slope Rule so that it can handle at least some transverse fiber products $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$ in which the manifolds **A**, **B**, and **S** are not all 1-dimensional.

2.2.1 Upward slopes versus downward slopes

What does the concept "Alice's terrain slopes upward" correspond to, in our new, multidimensional situation? The slope of Alice's terrain is one aspect of the local behavior of her altitude map $\mathbf{f}\colon \mathbf{A} \to \mathbf{S}$, say near the point $\mathbf{a}$ in $\mathbf{A}$. Using our standard abbreviations, the Jacobian of $\mathbf{f}$ at $\mathbf{a}$ is the matrix $[f]$, with a columns and s rows. In the special case that $s = a$, that Jacobian is square, so it makes sense to compute its determinant. When that determinant is positive, so that $\mathrm{sense}(\mathbf{f}, \mathbf{a}) = +1$, we say that Alice's terrain *slopes upward* at $\mathbf{a}$; when $\det([f]) < 0$, so that $\mathrm{sense}(\mathbf{f}, \mathbf{a}) = -1$, her terrain *slopes downward* at $\mathbf{a}$; when the Jacobian matrix $[f]$ is square but $\det([f]) = 0$, then her terrain is *flat*, in at least some direction — that is, there is some direction in which Alice can move without changing any of the s coordinates of her altitude to first order; and, when the Jacobian $[f]$ is not square, it doesn't make sense to talk about the slope of Alice's terrain.

In a similar way, it makes sense to talk about the slope of Bob's terrain only when $s = b$, so that the Jacobian $[g]$ of Bob's altitude map $\mathbf{g}$ will be square. Bob's terrain at $\mathbf{b}$ is then upward sloping, downward sloping, or flat according as $\det([g])$ is positive, negative, or zero.

2.2.2 Advancing versus retreating

What does the advice "Alice should advance" correspond to, in our new situation, where Alice has a dimensions to move around in? Consider Alice's projection map $\mathbf{u}\colon \mathbf{A} \times_{\mathbf{S}} \mathbf{B} \to \mathbf{A}$, the map defined by $\mathbf{u}(\mathbf{a}', \mathbf{b}') := \mathbf{a}'$. At the point $(\mathbf{a}, \mathbf{b})$ in the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$, the differential of the map $\mathbf{u}$ is the linear projection $u\colon A \times_S B \to A$. The Jacobian of $\mathbf{u}$ at $(\mathbf{a}, \mathbf{b})$ is the matrix $[u]$, which has $a+b-s$ columns (by transversality) and a rows. In the special case $s = b$, this Jacobian is square and might be invertible. When it is invertible, the map $\mathbf{u}$ provides a diffeomorphism between some neighborhood of $(\mathbf{a}, \mathbf{b})$ in the fiber product and some neighborhood of $\mathbf{a}$ in $\mathbf{A}$. We can use that diffeomorphism to compare some proposed orientation on $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ with the given orientation on $\mathbf{A}$. When we advise Alice to advance, we choose to orient the fiber product so that $\mathrm{sense}(\mathbf{u}, (\mathbf{a}, \mathbf{b})) = \mathrm{sgn}(u) = +1$, that is, so that $\mathbf{u}$ preserves orientation near $(\mathbf{a}, \mathbf{b})$; when we advise Alice to retreat, we choose the opposite orientation, the one that makes $\mathbf{u}$ reverse orientation there. (Warning: Once $a > 1$, we can no longer interpret Alice's advancing or retreating by thinking of Alice changing her position as a function of time. Time is inherently 1-dimensional — at least, in our universe! — while the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ has dimension $a + b - s = a$.)

In a similar way, it makes sense to advise Bob to advance or retreat only when $s = a$, so that the Jacobian $[v]$ of his projection map $\mathbf{v}\colon \mathbf{A} \times_{\mathbf{S}} \mathbf{B} \to \mathbf{B}$ is square, and only at a point $(\mathbf{a}, \mathbf{b})$ on the fiber product where that Jacobian $[v]$ is actually invertible. By advising Bob to advance or retreat, we are choosing our orientation on the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ so as to make the map $\mathbf{v}$ either preserve or reverse orientation at the point $(\mathbf{a}, \mathbf{b})$.

2.2.3 The Invertible Factor Laws

What do we get when we generalize the Partner's-Slope Rule? Let's first discuss how Alice controls her motion, based on the slope of Bob's terrain.

Note that we must have $s = b$ if we intend to measure Bob's slope as either upward or downward. Fortunately, that same condition is the one that is required if we intend to advise Alice to advance or retreat.

More precisely, we can measure Bob's slope as upward or downward at a point $(\mathbf{a}, \mathbf{b})$ on the fiber product just when Bob's altitude map $\mathbf{g}$ is a local diffeomorphism at $\mathbf{b}$. Let's suppose this to be the case. If we want to advise Alice to advance or retreat, Alice's projection map $\mathbf{u}$ must be a local diffeomorphism at $(\mathbf{a}, \mathbf{b})$. To see that this will also be the case, note that we can locally express the inverse of $\mathbf{u}$ in terms of the local inverse of $\mathbf{g}$ by setting $\mathbf{u}^{-1}(\mathbf{a}') := (\mathbf{a}', \mathbf{g}^{-1}(\mathbf{f}(\mathbf{a}')))$, where $\mathbf{a}'$ is close to $\mathbf{a}$ and hence $\mathbf{f}(\mathbf{a}')$ is close to $\mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b})$. In particular, the point $\mathbf{u}^{-1}(\mathbf{a}')$ lies on the fiber product because applying $\mathbf{f}$ to $\mathbf{a}'$ gives the same result as applying $\mathbf{g}$ to $\mathbf{g}^{-1}(\mathbf{f}(\mathbf{a}'))$. (Another way to prove this is to apply Exercise 2-1 to the differential maps in Figure 2.2, thereby deducing that the bijectivity of the differential $g = T_{\mathbf{b}}\mathbf{g}$ implies the bijectivity of the differential $u = T_{(\mathbf{a},\mathbf{b})}\mathbf{u}$.)

We thus arrive at the following law:

- If Bob's altitude map $\mathbf{g}$ is a local diffeomorphism at $\mathbf{b}$ that preserves orientation there (that is, if Bob's terrain slopes upward at $\mathbf{b}$), then Alice's projection map $\mathbf{u}$ will be a local diffeomorphism at $(\mathbf{a}, \mathbf{b})$ and we should orient the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ so that $\mathbf{u}$ preserves orientation there (that is, Alice should advance).

- Symmetrically, if $\mathbf{g}$ is a local diffeomorphism at $\mathbf{b}$ that reverses orientation (Bob's terrain slopes downward at $\mathbf{b}$), then $\mathbf{u}$ will again be a local diffeomorphism at $(\mathbf{a}, \mathbf{b})$, but we should orient the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ so that $\mathbf{u}$ reverses orientation there (Alice should retreat).

When Bob's altitude map $\mathbf{g}$ is not a local diffeomorphism at $\mathbf{b}$ — either because $s \neq b$, so that the Jacobian $[g]$ is not even square, or because the Jacobian is square but singular — then this law does not apply, and we gain no insight into the question of how to orient the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$. This law tells us how to orient the fiber product whenever the right-hand factor map $\mathbf{g}$ is locally invertible, so we shall refer to it as the *Right Invertible Law*.

The equation

$$\text{sense}(\mathbf{u}, (\mathbf{a}, \mathbf{b})) = \text{sense}(\mathbf{g}, \mathbf{b}) \tag{2-2}$$

is a more concise way to express the Right Invertible Law. Whenever the map $\mathbf{g}$ is a local diffeomorphism, and hence the right-hand side is defined, the map $\mathbf{u}$ will also be a local diffeomorphism and we should orient the fiber product — which is the domain of the map $\mathbf{u}$ — to make the left-hand side agree with the right.

The *Left Invertible Law* is symmetric, using Alice's slope to tell Bob whether to advance or retreat:

$$\text{sense}(\mathbf{v}, (\mathbf{a}, \mathbf{b})) = \text{sense}(\mathbf{f}, \mathbf{a}). \tag{2-3}$$

We'll refer to these two guidelines as the *Invertible Factor Laws*.

There is good news: Whenever both the Left and Right Invertible Laws apply, the advice that they give is always consistent. For both factor maps to be locally invertible, we must have $s = a$ and $s = b$, so we are in the equidimensional case. All four of the manifolds in Figure 2.1 then have the same dimension $a = b = s = a+b-s$. Furthermore, we must be at a point $(\mathbf{a}, \mathbf{b})$ at which both of the maps $\mathbf{f}$ and $\mathbf{g}$ are local diffeomorphisms. At such a point, both of the projection maps $\mathbf{u}$ and $\mathbf{v}$ are also local diffeomorphisms. The Left Invertible Law tells us to orient the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ so that $\text{sense}(\mathbf{v}, (\mathbf{a}, \mathbf{b})) = \text{sense}(\mathbf{f}, \mathbf{a})$, while the Right Invertible Law tells us to make $\text{sense}(\mathbf{u}, (\mathbf{a}, \mathbf{b})) = \text{sense}(\mathbf{g}, \mathbf{b})$. To see that those two pieces of advice are consistent, note that $\mathbf{f} \times_{\mathbf{S}} \mathbf{g} = \mathbf{f} \circ \mathbf{u} = \mathbf{g} \circ \mathbf{v}$ and hence we have $\text{sense}(\mathbf{f} \times_{\mathbf{S}} \mathbf{g}, (\mathbf{a}, \mathbf{b})) = \text{sense}(\mathbf{f}, \mathbf{a})\, \text{sense}(\mathbf{u}, (\mathbf{a}, \mathbf{b})) = \text{sense}(\mathbf{g}, \mathbf{b})\, \text{sense}(\mathbf{v}, (\mathbf{a}, \mathbf{b}))$. So our two pieces of advice boil down to the same thing, which we christen the *Both Invertible Law*:

$$\text{sense}(\mathbf{f} \times_{\mathbf{S}} \mathbf{g}, (\mathbf{a}, \mathbf{b})) = \text{sense}(\mathbf{f}, \mathbf{a})\, \text{sense}(\mathbf{g}, \mathbf{b}). \tag{2-4}$$

This simple and compelling identity tells us how to orient the fiber product at any point where both of the factor maps are locally invertible. Of course, at such a point, the Left Invertible Law by itself would also tell us what to do, as would the Right Invertible Law; but the Both Invertible Law tells us more elegantly, without our having to think about the projection maps $\mathbf{u}$ or $\mathbf{v}$.

But there is also bad news: Back in the 1-dimensional case of the original Alice and Bob, it followed from transversality that Alice and Bob could not be simultaneously at flat spots. Once the common dimension of the manifolds $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{S}$ exceeds 1, however, simultaneous flat spots can happen, even when the factor maps $\mathbf{f}$ and $\mathbf{g}$ are transverse. Indeed, there exist nonempty, transverse, equidimensional fiber products on which Alice and Bob are always at flat spots; see Exercise 7-1. No matter what point $(\mathbf{a}, \mathbf{b})$ we consider in such a fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$, the factor maps $\mathbf{f}$ and $\mathbf{g}$ are transverse at $(\mathbf{a}, \mathbf{b})$ and the Jacobians $[f]$ of $\mathbf{f}$ at $\mathbf{a}$ and $[g]$ of $\mathbf{g}$ at $\mathbf{b}$ are both square, but $\det([f]) = \det([g]) = 0$. The Invertible Factor Laws give us no advice at all about how to orient such a fiber product, no matter where on it we try to apply them.

By the way, since the Invertible Factor Laws are local, they make perfect sense also in the context of linear spaces and linear maps, as in Figure 2.2. Given such a fiber product of linear spaces, the Invertible Factor Laws tell us to set

$$\text{sgn}(u) = \text{sgn}(g) \tag{2-5}$$

$$\text{sgn}(v) = \text{sgn}(f) \tag{2-6}$$

$$\text{sgn}(f \times_S g) = \text{sgn}(f)\, \text{sgn}(g) \tag{2-7}$$

In each of these laws, if the maps on the right-hand side are invertible, so the signs on the right are well-defined, then the map on the left will also be invertible (as we saw in Exercise 2-1); and we should orient the fiber product $A \times_S B$, which is the domain of the left-hand map, so as to make the equality hold.

In summary, generalizing the Partner's-Slope Rule to higher dimensions leads to the Invertible Factor Laws. They specify how to orient many fiber products, and they never contradict each other. But they do not constitute a complete orientation rule, even for the equidimensional case.

Exercise 2-8 Exercise 1-1 discussed the Velocity Variant of the Partner's-Slope Rule in the 1-dimensional case. Explain why that variant does not generalize to the equidimensional case.

Answer: The core idea in Exercise 1-1 is for Alice to restrict the choice of the local coordinate system on the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ so as to make the Jacobian matrix $[u]$ of Alice's projection map $\mathbf{u}$ coincide with the Jacobian matrix $[g]$ of Bob's altitude map $\mathbf{g}$. Similarly, Bob restricts the choice to make his projection matrix $[v]$ coincide with Alice's altitude matrix $[f]$. Are those two restrictions consistent? The commutativity of Figure 2.2 tells us that $f \circ u = g \circ v$, from which we deduce the matrix equality $[f][u] = [g][v]$. Thus, we can arrange that $[u] = [g]$ and $[v] = [f]$ only when $[f][g] = [g][f]$ — that is, only when the Jacobian matrices of $\mathbf{f}$ and $\mathbf{g}$ commute. Of course, 1-by-1 matrices are simply scalars; they all commute. Once the common dimension of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{S}$ exceeds 1, however, we typically have $[f][g] \neq [g][f]$, so there won't be any local coordinate system on the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ that makes both $[u] = [g]$ and $[v] = [f]$. Thus, the idea that underlies the Velocity Variant typically isn't achievable.

2.3 Generalizing the Gray-Region Rule

The Gray-Region Rule generalizes more successfully than the Partner's-Slope Rule; but the process of generalization is a bit subtle. Indeed, as the first step in that process, I have a confession to make: The notion of the "gray region" doesn't always make sense, even in the 1-dimensional case — and hence the Gray-Region Rule, as we discussed it in Chapter 1, was a bit over-simplified.

2.3.1 The Gray-Side Rule

Consider a transverse fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ in which all three manifolds $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{S}$ are 1-dimensional and oriented, but the manifold $\mathbf{S}$ is a closed loop. Let $(\mathbf{a}, \mathbf{b})$ be a point in the direct product $\mathbf{A} \times \mathbf{B}$ at which $\mathbf{f}(\mathbf{a}) \neq \mathbf{g}(\mathbf{b})$; so $(\mathbf{a}, \mathbf{b})$ is not in the fiber product. Alice and Bob colored the state $(\mathbf{a}, \mathbf{b})$ either gray or white, according as they had $\mathbf{f}(\mathbf{a}) < \mathbf{g}(\mathbf{b})$ or $\mathbf{f}(\mathbf{a}) > \mathbf{g}(\mathbf{b})$ in $\mathbf{S}$. But the ability to discriminate between those two cases presupposes that the manifold $\mathbf{S}$ is totally ordered. If $\mathbf{S}$ is a closed loop, we can get from the point $\mathbf{f}(\mathbf{a})$ to the distinct point $\mathbf{g}(\mathbf{b})$ by moving either way around the loop, forward or backward. So it is no longer clear that there is such a thing as a “gray region”.

Indeed, when the manifold $\mathbf{S}$ is a loop in this way, removing the black path $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ from the state space $\mathbf{A} \times \mathbf{B}$ may fail to disconnect the latter manifold. The black path may behave like one of the circles that goes around a torus; removing such a circle from the torus leaves a cylinder, which is still connected. Thus, purely on topological grounds, there may be no hope of partitioning what is left into a “gray region” and a “white region”.

The fix for this difficulty is to work locally. Rather than defining the “gray region”, we define the *gray side* and the *white side* of the black path locally by looking at what tiny motions of Alice or Bob would do to their relative altitudes. Indeed, this is what we actually did; what Figure 1.10 actually shows, for example, are the gray and white sides of a tiny segment of the black path. Alice and Bob can figure out which direction to walk once they know which side of the black path is gray locally; they have no need even to think about the entire gray region. We refer to this clarification of the Gray-Region Rule as the *Gray-Side Rule*.

Note that, given a circle that goes around a torus, we can make a consistent convention about the gray side and the white side of that circle, even though we cannot partition the cylinder that remains after removing the circle into a “gray region” and a “white region”. Note also that the open arrowheads generated by the Gray-Side Rule will always be locally consistent, even in cases where no “gray region” can be globally defined. To verify this local consistency, it suffices to check that nearby applications of the Gray-Side Rule will agree about which side of the black path is the gray side.

2.3.2 Orienting direct products and quotients

At this point, we need some observations about orienting the direct products and quotients of linear spaces. The key point is that arbitrary choices are involved.

Given oriented linear spaces V and W, we can always put an orientation on their direct product $V \times W$. We typically do so by concatenating a positive basis for V and a positive basis for W, with the basis for V going first, and specifying that the resulting concatenation constitutes a positive basis for $V \times W$. But choosing V to go first in this concatenation is arbitrary, since the direct product $V \times W$ is isomorphic to $W \times V$. For example, suppose that V and W are both copies of the real numbers $\mathbb{R}$, each oriented so that $(+1)$ is a positive basis. Their direct product $V \times W$ is a plane. We can orient that plane so that rotation from the positive V axis to the positive W axis is either clockwise or counterclockwise. The choice is arbitrary, as we saw by contrasting Figure 1.7 with Figure 1.11.

A similar issue arises for quotients. Suppose that U is a linear subspace of a linear space V. Given any basis for U, we can extend that basis into a basis for V. And the vectors that we add, in doing that extension — more precisely, their equivalence classes modulo U — form a basis for the quotient space V/U. If we choose the basis for U to be positive and we extend it so as to form a positive basis for all of V, we can orient the quotient V/U by specifying that the vectors added during this extension constitute a positive basis for V/U. But this construction also involved an arbitrary choice: We chose to put the basis for U first and the basis for V/U second, in assembling the basis for V, when the other order would have been just as good. Thus, in orienting the quotient V/U, we are forced to make an arbitrary choice between $V \cong U \times (V/U)$ and $V \cong (V/U) \times U$.

Such quotients often arise from linear maps. Let $m\colon V \to W$ be a linear map between linear spaces, let $\operatorname{Dom}(m) := V$ be the domain of m, let $\operatorname{Ker}(m) := m^{-1}(\{0\})$ be the kernel of m in V, and let $\operatorname{Im}(m) := m(V)$ be the image of m in W. It is a basic theorem of linear algebra that the image is canonically isomorphic to the domain modulo the kernel: $\operatorname{Im}(m) \cong \operatorname{Dom}(m)/\operatorname{Ker}(m)$. Hence, given orientations on any two of those three spaces, we can orient the third. But doing so also involves making an arbitrary choice; roughly speaking, we must choose between realizing $\operatorname{Dom}(m)$ as $\operatorname{Ker}(m) \times \operatorname{Im}(m)$ or as $\operatorname{Im}(m) \times \operatorname{Ker}(m)$.

2.3.3 The Delta Rule

The Gray-Side Rule generalizes quite successfully into a rule that handles the full, any-dimensional case. But doing so involves making arbitrary choices, which will come back to haunt us.

Consider some any-dimensional, transverse fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$, and consider the neighborhood of some point $(\mathbf{a}, \mathbf{b})$ in $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$. We adopt our standard abbreviations: So the point $\mathbf{s}$ in $\mathbf{S}$ is the common value $\mathbf{s} := \mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b})$. We denote by A the tangent space $A := T_{\mathbf{a}}\mathbf{A}$ and similarly for $B := T_{\mathbf{b}}\mathbf{B}$ and $S := T_{\mathbf{s}}\mathbf{S}$. The differential of $\mathbf{f}$ at $\mathbf{a}$ is the linear map $T_{\mathbf{a}}\mathbf{f}\colon T_{\mathbf{a}}\mathbf{A} \to T_{\mathbf{s}}\mathbf{S}$, which we abbreviate as $f\colon A \to S$. And similarly for $g\colon B \to S$. The linear maps f and g are themselves transverse, and the tangent space to the fiber product is the fiber product of the tangent spaces: $T_{(\mathbf{a},\mathbf{b})}(\mathbf{A} \times_{\mathbf{S}} \mathbf{B}) \cong A \times_S B$. Our challenge is to orient this tangent space $A \times_S B$, given orientations on the tangent spaces A, B, and S.

Note that this challenge, because it is local, involves only linear spaces and linear maps. To meet the challenge, we define a linear map $\Delta\colon A \times B \to S$ by subtracting, setting

$$\Delta(\alpha, \beta) := g(\beta) - f(\alpha),$$

for any point (α, β) in the direct product $A \times B$. If Alice moves from the point $\mathbf{a}$ to some nearby point $\mathbf{a}'$, a certain vector α describes, to first order, how she has moved. Similarly, a certain vector β describes how Bob has moved from $\mathbf{b}$ to $\mathbf{b}'$. If Alice and Bob both move in this way, the vector $\Delta(\alpha, \beta)$ tells us, to first order, the discrepancy that arises between the points $\mathbf{f}(\mathbf{a}')$ and $\mathbf{g}(\mathbf{b}')$. The kernel of Δ is the space of all directions in which Alice and Bob can move away from $(\mathbf{a}, \mathbf{b})$ while preserving the relationship $\mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b})$ to first order. Thus, that kernel is precisely the linear-space fiber product $\operatorname{Ker}(\Delta) = A \times_S B$ that we want to orient.

This map Δ is the key to our generalized Gray-Side Rule. We are given orientations on the manifolds $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{S}$ and hence on their tangent spaces A,

B, and S. By concatenation, we can put an orientation on the direct product $A \times B = \mathrm{Dom}(\Delta)$. Because we are restricting our attention to transverse fiber products, it turns out that $\mathrm{Im}(\Delta)$ is the entire tangent space $\mathrm{Im}(\Delta) = S$, for which we have a given orientation. Putting together those orientations on $\mathrm{Dom}(\Delta)$ and $\mathrm{Im}(\Delta)$, we get an orientation on $\mathrm{Ker}(\Delta) = A \times_S B$, which was our goal. We refer to this orientation technique as the *Linear-Space Delta Rule*.

The *Smooth-Manifold Delta Rule* works in the obvious way; it orients each tangent space $A \times_S B$ to the fiber-product manifold $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ by applying the Linear-Space Delta Rule to that tangent space. The resulting orientations on the tangent spaces always fit together in a locally consistent manner, thereby orienting the manifold $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ itself; we prove this local consistency in Section 7.3 by showing that the Linear-Space Delta Rule has the continuity property called *stability*, which we define in Section 6.2.

So the good news is that the Smooth-Manifold Delta Rule is powerful enough to orient all transverse fiber products, even the any-dimensional ones. The bad news is that we are forced to make various arbitrary choices of convention in implementing the Delta Rule. Which basis goes first, A or B, in concatenating a basis for $A \times B$? Which goes first, $\mathrm{Ker}(\Delta)$ or $\mathrm{Im}(\Delta)$, in concatenating a basis for $\mathrm{Dom}(\Delta)$? For that matter, why did we set $\Delta(\alpha, \beta) := g(\beta) - f(\alpha)$, rather than $\Delta(\alpha, \beta) := f(\alpha) - g(\beta)$? Because of those arbitrary choices, we have no guarantee that the orientations produced by any particular variant of the Delta Rule will be natural and hence have good mathematical properties. Instead, the orientations that we get will depend upon the arbitrary choices that we have made. Thus, we must add an explicit fudge factor to the Delta Rule in order to "calibrate" it, that is, to arrange that it gets the "proper" orientations as its answers.

2.4 Defining the Proper Orientations

But which orientations of transverse fiber products are "proper"? That question turns out to be quite easy in the 1-dimensional case, fairly easy also in the equidimensional case, but rather subtle in the any-dimensional case. All three of those cases arise in two different problem domains: linear spaces and smooth manifolds. But any orientation rule for smooth manifolds that is local reduces to an rule for linear spaces; so the linear-space problem is the fundamental one.

2.4.1 Linear-space orientation rules as subroutines

Our overarching goal is an orientation rule for transverse fiber products of smooth manifolds. But we have decided up front that this rule must be local; that is, it must operate by using some orientation rule for transverse fiber products of linear spaces as a subroutine and applying that subroutine to each tangent space of the manifold independently. The answers given back from the smooth-manifold rule are completely determined by the answers from its linear-space subroutine. So, in considering issues of "propriety", it suffices to focus on the subroutine, that is, to consider orientation rules for transverse fiber products of linear spaces.

Not any linear-space rule will do, of course. The orientations on the various tangent spaces returned by the linear-space subroutine must be locally consistent, so that they fit together to give an orientation on the manifold. To guarantee that all of our linear-space rules behave well in this regard, we require that they be *stable*, in the sense defined in Section 6.2. We prove in Section 6.3 that any stable linear-space rule will always orient the tangent spaces in a locally consistent manner and will hence lift to give a smooth-manifold orientation rule. Requiring stability thus addresses the issue of local consistency, leaving us free to focus on the linear-space question: Given a transverse fiber product of linear spaces, which orientation of that fiber product is "proper"?

2.4.2 The easy cases

In the 1-dimensional case of Alice and Bob, choosing the proper orientation is quite easy. When both Alice and Bob are on upward slopes, we were never in any doubt that they should both advance; so the *proper* open arrowhead in the upper-left cell in Figure 1.10 points toward the upper right. The other seven cases are then determined by stability: As the black line rotates, the open arrowhead must rotate with it, in order to avoid any sudden reversals.

The equidimensional case turns out to be almost as easy. We consider first an equidimensional fiber product $A[f] \times_S [g]B$ in which both of the factor maps f and g are invertible. The Both Invertible Law 2-7 then convincingly states that the *proper orientation* on the fiber product is the one that makes $\operatorname{sgn}(f \times_S g) = \operatorname{sgn}(f)\operatorname{sgn}(g)$. We show in Section 7.5 that all other equidimensional cases are determined from this by stability.

We could use a similar argument, based on stability and the Left Invertible Law, to define the proper orientation for fiber products in which $a = s \neq b$. And the Right Invertible Law would handle the ones in which $a \neq s = b$. But the full, any-dimensional case is quite a different story. When $a \neq s \neq b$, so that neither factor map can possibly be invertible, it isn't at all clear which of the two possible orientations is the proper one.

2.4.3 Axioms for the proper orientations

The way to investigate that question is to invent axioms that constrain the behavior of a linear-space orientation rule for the any-dimensional case. Indeed, each of the Invertible Factor Laws is essentially such an axiom: the Identity Axioms in Sections 9.1.4 through 9.1.6. We succeed in this axiomatic strategy if we can invent a compelling collection of axioms that is both consistent and complete. "Consistent" here means that there is some orientation rule that satisfies all of our axioms. "Complete" means that there is at most one such rule. So, if we find an axiom system that is both consistent and complete, there will be a unique orientation rule that satisfies all of our axioms. We can then define the *Proper Orientations* to be the orientations produced by that unique rule. And we can calibrate the Delta Rule so that it produces those Proper Orientations.

Associativity is one obvious axiom to try for. Let A, B, C, and S be oriented linear spaces and let $f: A \to S$, $g: B \to S$, and $h: C \to S$ be linear maps that are transverse in the appropriate sense. The two linear spaces $(A \times_S B) \times_S C$ and $A \times_S (B \times_S C)$ are then canonically isomorphic, and we surely want to require that our rule orients them so that the canonical isomorphism preserves orientation. That is, we want to have $(A \times_S B) \times_S C = A \times_S (B \times_S C)$, rather than having $(A \times_S B) \times_S C = -\big(A \times_S (B \times_S C)\big)$. Unfortunately, all of the obvious axioms, including associativity, are not enough to narrow down the space of orientation rules for the any-dimensional case to a single, proper rule.

The key to finding an axiom system that is complete is to insist on a stronger form of associativity:

$$(A \times_S B) \times_T C = A \times_S (B \times_T C).$$

In this *Axiom of Mixed Associativity*, the two fiber products are over different base spaces S and T. The implicit maps that are really being multiplied, in such a mixed fiber product, have the zigzag form:

$$\begin{tikzcd} A \arrow[dr, "f"'] & & B \arrow[dl, "g"] \arrow[dr, "h"'] & & C \arrow[dl, "k"] \\ & S & & T & \end{tikzcd}$$

Note that this zigzag involves two factor maps from B, a "forward" one to T and a "backward" one to S. By adding this Axiom of Mixed Associativity to our more straightforward axioms in Chapter 9, we produce a collection of axioms that is both consistent and complete, thereby defining the Proper Orientations for the any-dimensional case.

2.5 Overview

Now is perhaps a good time for an overview of this monograph.

In Chapter 1, we familiarized ourselves with fiber products by climbing mountains with Alice and Bob. We came up with two rules for orienting 1-dimensional fiber products: the Partner's-Slope Rule and the Gray-Region Rule.

Our overall goal is to orient the smooth manifolds that result when we take transverse fiber products of oriented smooth manifolds. We want to do so, not only in the 1-dimensional case of Alice and Bob, but also in the equidimensional and any-dimensional cases. In the current Chapter 2, we have considered the extent to which the Partner's-Slope Rule and the Gray-Region Rule can be generalized. The Partner's-Slope Rule generalizes into the Invertible Factor Laws, which are essentially axioms about how an orientation rule for fiber products should behave.

Those axioms do specify the Proper Orientation of many fiber products — but not all of them, not even all of the equidimensional ones. The Gray-Region Rule generalizes into the Delta Rule, a framework that is powerful enough to handle the full any-dimensional case. Unfortunately, building this framework involves making some arbitrary choices. So the resulting Delta Rule needs to be calibrated with an explicit fudge factor, before it will assign the Proper Orientations.

Chapter 3 completes the introduction by providing some motivation for all this. We discuss a problem of practical interest, arising in CAGD and robotics, where it is important that the fiber products of oriented manifolds themselves be oriented. In particular, we discuss computing the Minkowski sums and convolutions of regions that are described by specifying their boundaries.

Our mathematical work begins in earnest in Chapter 4, where we study the direct products and fiber products of sets, linear spaces, and smooth manifolds. For linear spaces and smooth manifolds, where the concept of dimension makes sense, we discuss the standard notion of transversality.

In Chapter 5, we study what it means to orient a linear space or a smooth manifold. While our definitions here are also essentially standard, we must take care to arrange that all linear spaces end up with two possible orientations, even linear spaces of dimension zero (such a space having the empty sequence of vectors as its only basis). We discuss how orientation interacts with direct products and with quotient spaces. Given a linear map between two spaces that are each direct products, we discuss how to represent that linear map as a matrix whose entries are themselves linear maps. We can perform elementary row and column operations on the resulting matrices, just as if they were matrices of numbers; but we must be careful to “multiply” by composing from the correct side.

Stability is the key concept in Chapter 6. We want our orientation rule for smooth manifolds to be local, meaning that it must operate by using an orientation rule for linear spaces as a subroutine and applying that subroutine to each tangent space independently. For this to work, the linear-space rule must have a certain continuity property that we christen *stability*. In Chapter 6, we define stability, and we prove that any stable linear-space rule will orient the various tangent spaces of a smooth manifold in a locally consistent manner, hence giving us an orientation rule for smooth manifolds. In the remainder of the monograph, we can hence focus on stable orientation rules for linear-space fiber products.

Chapter 7 presents our initial, uncalibrated version of the Delta Rule for linear spaces and proves that it is stable. We use the Both Invertible Law and stability to define the Proper Orientation for any transverse, equidimensional fiber product, and we demonstrate that even the Uncalibrated Delta Rule assigns the Proper Orientations in the equidimensional case. Recall that the Left and Right Invertible Laws apply also to some transverse fiber products that are not equidimensional. We find that the Uncalibrated Delta Rule assigns orientations to some of those cases that violate the Left Invertible Law. This is convincing evidence that the Delta Rule must be explicitly calibrated in order to get the Proper Orientations, once we leave the equidimensional case.

The key to defining the Proper Orientations in the any-dimensional case is the Axiom of Mixed Associativity. Chapter 8 returns to the basic definitions from Chapter 4, generalizing them so that they can handle mixed fiber products, where a different base space is involved for each adjacent pair of factor spaces. In the process, we define a *zigzag* to be a particular structure in linear algebra: a sequence of linear spaces in which adjacent spaces are related by linear maps that alternate in direction.

Everything comes to a head in the Chapter 9. We present a list of axioms that contains the Axiom of Mixed Associativity. We demonstrate that those axioms are both consistent and complete, so the orientations that they describe are the unique Proper Orientations. We then calculate the fudge factor needed to calibrate the Delta Rule so that it produces those Proper Orientations.

In Chapter 10, we discuss the freedom that we would acquire to adopt other orientation rules if we were to abandon certain of our axioms. We also discuss several formulas that bundle together subsets of our axioms. The fanciest of these is the *Binary Full Formula*,

$$(L \times N \times Q) \times_{(M \times N \times Q)} (M \times N \times R) = L \times N \times R,$$

a single identity that is free of fudge factors and yet is powerful enough, all by itself, to capture the Proper Orientation for every transverse fiber product of linear spaces. The subtle point about the Binary Full Formula is the order of the factors in each of its ternary direct products.

As part of understanding why the Binary Full Formula is compatible with mixed associativity, we also write out a Ternary Full Formula in Chapter 10. And we show that it would be possible to write out an n-ary Full Formula, for any n. Unfortunately, these n-ary Full Formulas for $n \geq 3$ are unavoidably asymmetric, and hence are not as pretty as the Binary Full Formula. To see why these higher full formulas have the universality implied by the adjective "full", we need to appeal to an intriguing theorem of linear algebra about the structure of zigzags — a theorem that follows, fortunately, from Gabriel's Theorem in the theory of quiver representations.

Chapter 3

Minkowski sum and convolution

Recall that our goal in this monograph is to orient the smooth manifolds that result when we take transverse fiber products of oriented smooth manifolds. In this third and final introductory chapter, we motivate that problem by discussing an application area in which it arises: *Minkowski sums* and *convolutions* in CAGD and robotics. For more details about these issues, see [1, 6, 13].

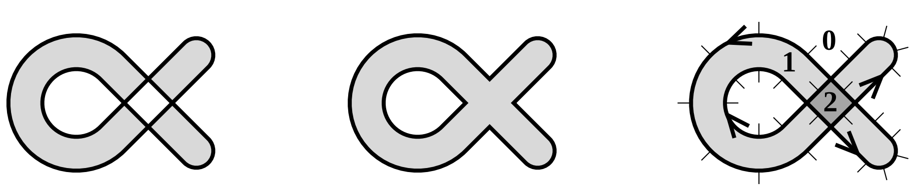
3.1 The oriented boundaries of paintings

It often works well to represent a region in d-space by specifying its boundary, since that boundary has dimension only $(d-1)$. For this reason, representation schemes based on boundaries are common in CAGD and robotics.

In such schemes, the boundaries are typically taken to be oriented. Perhaps the clearest justification for that decision comes from examples like Figure 3.1. Think of the planar region in that figure as the letter alpha from a Greek font (drawn by a type designer who is far too enamored of circles and straight lines).

In Figure 3.1(a), the boundary of the alpha is an unoriented curve. Without an orientation to provide guidance, the only local rule for distinguishing the inside from the outside asserts that every crossing of the boundary takes us either from inside to outside or vice versa. By that rule, unfortunately, the diamond where the alpha crosses itself ends up being outside — which would never do.

We could restructure the boundary as shown in Figure 3.1(b). Note that the boundary now consists of two separate loops, the outer of which has three sharp



(a) (b) (c)

Figure 3.1: A letter α, drawn three different ways

corners, while the inner has one. But it may be expensive to compute the four self-intersections and to restructure the boundary in that way. Furthermore, it would be delicate to edit this two-loop version of the boundary while preserving the illusion of a single brush stroke that crosses itself.

Orienting the boundary, as shown in Figure 3.1(c), provides a better solution. We then make the convention that crossing the boundary, say, from its right to its left, means moving from outside to inside. Under this convention, the diamond of self-intersection lies inside the alpha twice. This makes some intuitive sense, since the points in that diamond would be passed over twice as a brush painted the alpha. In computer graphics and in complex analysis, this situation is described by saying that the *winding number* of the oriented boundary around points in that diamond is 2.

Once some points are to lie inside the alpha with multiplicity 2, we can no longer view the alpha simply as a region, that is, as a subset of the plane. Rather, it is at least a multiset. In fact, we want to allow our winding numbers to be negative, as well as positive, since reversing the orientation on the boundary should negate all of the resulting winding numbers. So the alpha has become a function from the plane to the integers whose regions of constancy are well-behaved, in some sense yet to be specified. That concept reminds me of painting by number, so let's refer to such an integer-valued function as a *painting*. The alpha in Figure 3.1(c) is a painting in the plane that takes on the values 0, 1, and 2, as encoded by the colors white, light gray, and dark gray.

Warning: While we use the concept of a "painting" in this chapter, we don't define that concept precisely. In particular, we don't specify how well-behaved the regions of constancy have to be. That sloppiness is permissible because this chapter provides only motivation, not mathematics; paintings are just one example of a practical situation that needs oriented fiber products. The best candidate that I know of for a precise notion of "painting" is Schapira's notion of a *constructible function* [15]. Under Schapira's definition, the regions of constancy of a painting are the strata of a subanalytic stratification.

One final remark about the graphical conventions used in Figure 3.1(c): The standard way to draw an orientation on a smooth manifold M involves some sort of tangential graphical structure, something that indicates an orientation on sample tangent spaces. For example, we draw an orientation on a 1-manifold by drawing arrowheads pointing along it, forward or backward. When the manifold M sits, as a submanifold, inside an oriented manifold of some larger dimension, we have the option of drawing some sort of normal structure instead, something that indicates an orientation on sample complements of tangent spaces. For example, when M is a 1-dimensional submanifold of an oriented 2-manifold, we can indicate our choice of orientation on M by drawing, say, outward-pointing unit normal vectors along M. In Figure 3.1(c), we have redundantly used both schemes, both arrowheads along M and short, headless, outward-pointing vectors normal to M. We refer to the latter vectors as *whiskers*. We introduce these two different schemes for indicating the orientation of M because, in later examples, we shall be deal-

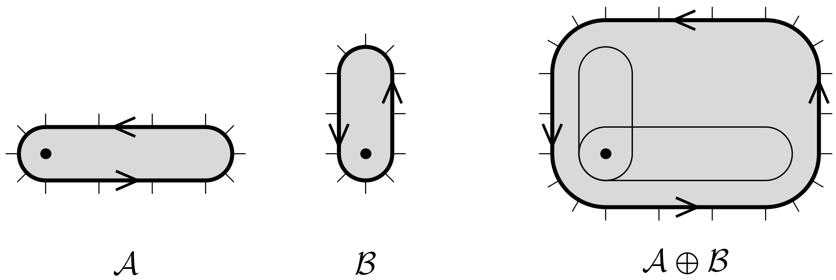


Figure 3.2: The Minkowski sum of two orthogonal rods

ing with two orientations that don't always agree. In this example, however, the whiskers and the arrowheads always do agree; that is, the whiskers always point toward the right-hand barb[1] of the arrowhead.

3.2 Minkowski sums of convex regions

Given two regions $\mathcal{A}$ and $\mathcal{B}$ in the plane, their *Minkowski sum* is the region

$$\mathcal{A} \oplus \mathcal{B} := \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in \mathcal{A} \text{ and } \mathbf{b} \in \mathcal{B}\},$$

where the plus in $\mathbf{a} + \mathbf{b}$ denotes vector sum. For example, the Minkowski sum of the horizontal rod $\mathcal{A}$ and the vertical rod $\mathcal{B}$ in Figure 3.2 is the rounded rectangle $\mathcal{A} \oplus \mathcal{B}$ on the right. Its boundary consists of four line segments, each of which is a translate of a flat side of one of the two rods, together with four quarter circles, each with twice the radius of the semicircles at the ends of the rods.

Given the boundaries of the two regions $\mathcal{A}$ and $\mathcal{B}$, how can we compute the boundary of their Minkowski sum $\mathcal{A} \oplus \mathcal{B}$?

If we start with a point $\mathbf{a}$ in the interior of $\mathcal{A}$, then every point of the form $\mathbf{a} + \mathbf{b}$, for $\mathbf{b}$ in $\mathcal{B}$, will lie in the interior of the Minkowski sum. Thus, the only way to get out to the boundary of the Minkowski sum $\mathcal{A} \oplus \mathcal{B}$ is to add a point $\mathbf{a}$ that lies on the boundary of $\mathcal{A}$ to a point $\mathbf{b}$ that lies on the boundary of $\mathcal{B}$.

But we need more. If the tangent line to the region $\mathcal{A}$ at $\mathbf{a}$ is not parallel to the tangent to $\mathcal{B}$ at $\mathbf{b}$, then we can combine tiny motions along the boundaries of $\mathcal{A}$ and $\mathcal{B}$ to nudge the sum $\mathbf{a} + \mathbf{b}$ an infinitesimal distance in any direction; so, once again, we can't possibly be out on the boundary of the Minkowski sum. To get out to the boundary of the sum $\mathcal{A} \oplus \mathcal{B}$, we must add points $\mathbf{a}$ and $\mathbf{b}$ on the boundaries of $\mathcal{A}$ and $\mathcal{B}$ where the tangent lines are parallel.

Even requiring the tangents to be parallel is not restrictive enough. Consider all of the boundary points of the regions $\mathcal{A}$ and $\mathcal{B}$ in Figure 3.2 at which the tangent is horizontal. Each point along the top edge of the Minkowski sum $\mathcal{A} \oplus \mathcal{B}$

[1] By the "right-hand barb", we mean the barb on your right when you face in the direction that the arrow points; the nautical term would be "starboard barb".

is the vector sum of a point along the top edge of $\mathcal{A}$ and the unique point at the middle of $\mathcal{B}$'s top semicircle; and the same is true if we replace "top" throughout with "bottom". But adding a top boundary point of $\mathcal{A}$ to the middle of $\mathcal{B}$'s bottom semicircle leads to a point of $\mathcal{A} \oplus \mathcal{B}$ that is not on the boundary. Thus, all of the points on the boundary of the Minkowski sum $\mathcal{A} \oplus \mathcal{B}$ have the form $\mathbf{a} + \mathbf{b}$ where $\mathbf{a}$ and $\mathbf{b}$ are points on the boundaries of $\mathcal{A}$ and $\mathcal{B}$ at which the tangent lines match both in slope and in orientation.

Beware! We are now starting to use the orientations on our boundaries for two different purposes. We continue to use the orientation on the boundary of any painting to compute the values of that painting, that is, to determine which side of the boundary has the larger winding number. But now, when forming the boundary of a Minkowski sum $\mathcal{A} \oplus \mathcal{B}$, we also compare the orientations on the boundaries of $\mathcal{A}$ and $\mathcal{B}$ in order to decide which of the tangents of $\mathcal{A}$ should be added to which of the tangents of $\mathcal{B}$. To avoid confusion in what follows, let's make an association between these two different uses of the orientation and our two different graphical conventions for drawing that orientation. We'll use the arrowheads that point along the boundary to compute winding numbers, under the convention that crossing a boundary from its right to its left increases the winding number by 1. When matching up tangents to compute the boundary of a Minkowski sum $\mathcal{A} \oplus \mathcal{B}$, we'll use the whiskers instead; we'll look for pairs of points $\mathbf{a}$ and $\mathbf{b}$ on the boundaries of $\mathcal{A}$ and $\mathcal{B}$ at which the whiskers are equal as vectors — not only parallel, but also with the same sense.

We now have a glimmer of why fiber products arise in computing Minkowski sums. Let $\mathcal{A}$ and $\mathcal{B}$ be two regions in the plane whose boundaries A and B are smooth 1-manifolds. Let $\mathbb{S}^1$ denote the unit circle, the set of all unit-length vectors in the plane. We can map A to $\mathbb{S}^1$ by taking each boundary point to the whisker at that point; call that map $w_A\colon A \to \mathbb{S}^1$ the *whisker map* (or Gauss map), and define the whisker map $w_B\colon B \to \mathbb{S}^1$ in a similar way. The boundary of the Minkowski sum $\mathcal{A} \oplus \mathcal{B}$ then corresponds to $A[w_A] \times_{\mathbb{S}^1} [w_B]B$, the fiber product of the boundaries of the two summands taken over the circle $\mathbb{S}^1$, where the factor maps from A and B to $\mathbb{S}^1$ are the whisker maps.

But we don't yet have any clear motivation for orienting the fiber products of oriented manifolds. While the boundaries A, B, and $A \times_{\mathbb{S}^1} B$ of the regions $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{A} \oplus \mathcal{B}$ are in fact oriented manifolds, there is no reason as yet why we need to keep track of that orientation as a separate concept. Since we know the whisker maps, we can compute the correct arrowhead at any point on any of these boundaries by rotating the whisker vector 90 degrees counterclockwise.

Exercise 3-1 Given two regions $\mathcal{A}$ and $\mathcal{B}$ in the plane with smooth boundaries A and B, suppose that we want the whisker maps $w_A\colon A \to \mathbb{S}^1$ and $w_B\colon B \to \mathbb{S}^1$ to be transverse. What geometric condition must we impose?

Answer: The analog of a plateau on Alice's mountain range is a segment of the boundary A that is straight; along such a segment, the whisker map is constant. The analog of a peak or a valley is a point on A where the boundary switches

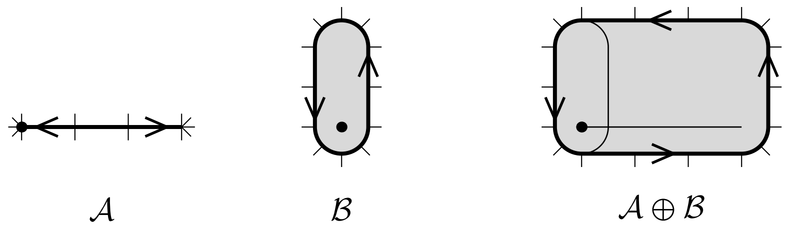


Figure 3.3: Sliding the brush $\mathcal{B}$ along the trajectory $\mathcal{A}$

between curving to the left and curving to the right. Such points are called *points of inflection* or *flexes*; at a flex, the whisker map has a local extremum. Technically, the points along a straight segment also count as flexes. So, if we want the whisker maps w_A and w_B to be transverse, what we must avoid is having flexes, one on A and the other on B, whose whiskers are equal as vectors — that is, are parallel and have the same sense.

3.3 1-dimensional regions as trajectories

By the way, there was no reason why the two rods in Figure 3.2 had to be the same thickness. Figure 3.3 shows the degenerate situation in which the rod $\mathcal{A}$ has shrunk to have zero width. Thus, the upper and lower boundaries of $\mathcal{A}$ run along, one right on top of the other.

When one of the summands of a Minkowski sum is 1-dimensional in this sense, we can reinterpret the summation as a dynamic process. We view the 1-dimensional summand as a trajectory and the other summand as a brush, and we translate the brush along the trajectory. In Figure 3.3, we get the stroke $\mathcal{A} \oplus \mathcal{B}$ by translating the rod-shaped brush $\mathcal{B}$ along the trajectory $\mathcal{A}$ (which happens to be straight, but wouldn't have to be). Indeed, Minkowski sums are used in graphics to model brush strokes in precisely this way. But note that the motion of the brush must be limited to pure translation. If the brush were to rotate or to change shape as it moved along the trajectory, then we would be dealing, not with a Minkowski sum, but with a more general problem in differential geometry.

When two segments of boundary run along right on top of each other, as do the top and bottom boundaries of the region $\mathcal{A}$ in Figure 3.3, our conventions for drawing orientations can become ambiguous. In particular, there's nothing in the picture to indicate that the rightward-pointing arrowhead along $\mathcal{A}$ goes with the downward-pointing whiskers, while the leftward-pointing arrowhead goes with the upward-pointing whiskers. We trust to context to resolve such ambiguities.

Minkowski sums of higher dimension arise in robotics. For example, consider translating a sofa around in a living room — but only translating, not rotating. If we Minkowski-subtract the sofa from the living room, we are left with the set of collision-free positions for the sofa under translation.

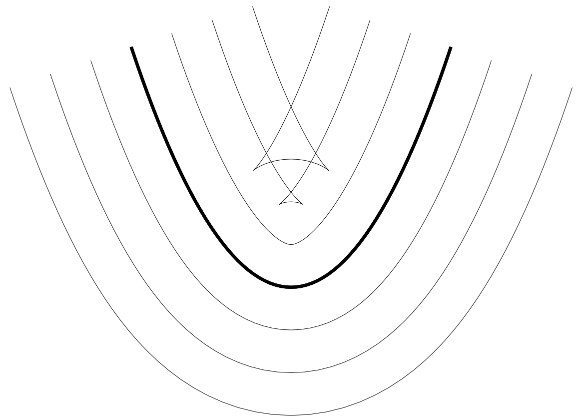


Figure 3.4: Offsets of a parabolic arc

3.4 Concave boundaries with tight turns

To prepare for our next example of a Minkowski sum, let's look at some offsets of a parabola. The darker curve in Figure 3.4, the middle one of the seven, is that portion of the graph of the standard parabola $y = x^2$ that lies over the x interval $[-3/2 \,..\, 3/2]$. The three curves on either side are the curves that result from offsetting that parabola by a distance of $\pm 2/5$, $\pm 4/5$, and $\pm 6/5$. We could imagine drawing Figure 3.4 by building a beam with seven pens, spaced $2/5$ apart, and then sliding that beam along the parabola. We keep the beam centered on the parabola and normal to it at all times — like a tightrope walker carrying a beam for balance. The surprising features of the result are the swallowtails that appear in the paths traced by the two innermost pens. A swallowtail results when the forward motion of a pen that arises from our progress along the parabola is overwhelmed by the backward motion that arises from our rotation of the beam. The farther a pen sticks out toward the inside of the turn, the more likely that pen is to experience such retrograde motion.

Exercise 3-2 Let r be a distance large enough so that the inner offset to the parabola at distance r has a swallowtail. (For the standard parabola $y = x^2$, that means $r > 1/2$.) Where are the two cusps of that swallowtail located?

Answer: The two cusps are the centers of the two circles of radius r that osculate the parabola, that is, that match it in position, slope, and curvature.

With swallowtails in mind, consider the Minkowski sum in Figure 3.5. The region $\mathcal{A}$ is the same parabolic arc that we saw in Figure 3.4, but now viewed as a 1-dimensional region. Since $\mathcal{A}$ is 1-dimensional, we can interpret the sum $\mathcal{A} \oplus \mathcal{B}$ as translating the brush $\mathcal{B}$ along the trajectory $\mathcal{A}$, where the brush $\mathcal{B}$ is a circular disk of radius $6/5$. If we compute the boundary of the stroke $\mathcal{A} \oplus \mathcal{B}$ by adding pairs of boundary points with equal whiskers, we get the painting shown in Figure 3.5 — complete with the swallowtail in its upper boundary.

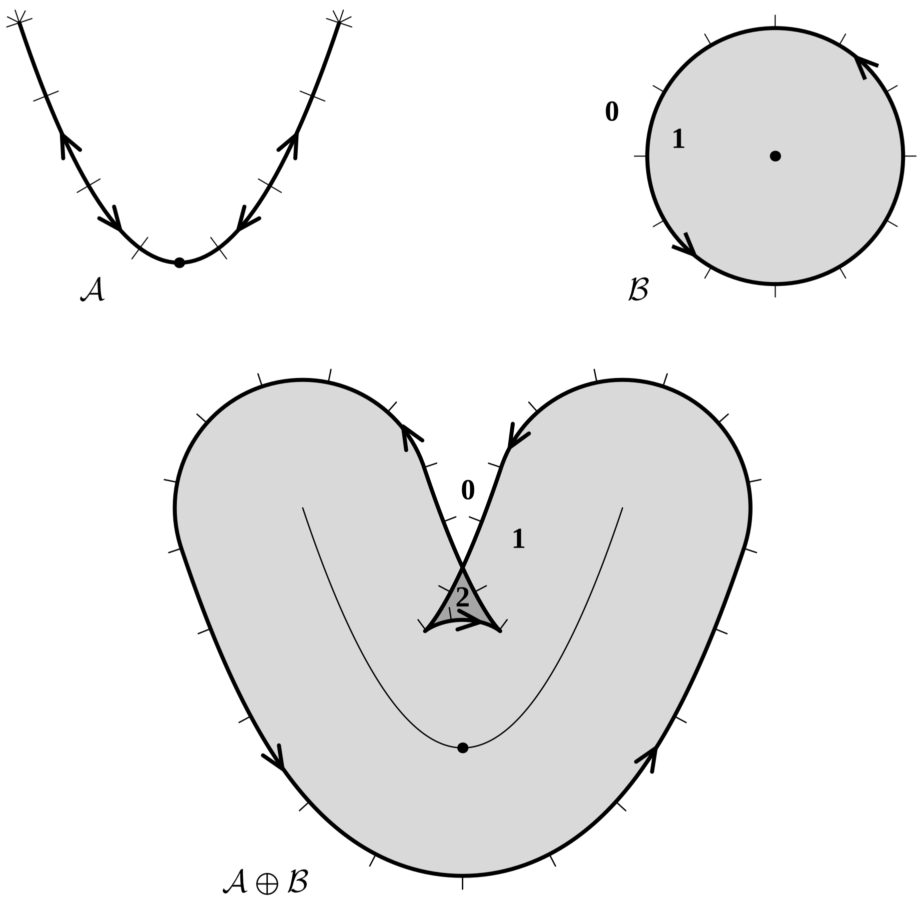


Figure 3.5: A parabolic trajectory with a tight turn

Such swallowtails can't arise in Minkowski sums whose summands are both convex, like those shown in Figures 3.2 and 3.3. The whisker map of a convex region is weakly monotonic; that is, advancing along that region's boundary may cause the whisker to rotate counterclockwise or may leave the whisker's azimuth unchanged, but it never causes the whisker to rotate clockwise. Borrowing some terminology from Alice and Bob, all of the terrain on such a mountain range is either upward-sloping or flat; none of it slopes downward. If both mountain ranges have this character, then Alice and Bob can traverse the fiber product without either ever needing to retreat, so no retrograde motion arises.

But retrograde motion can arise in a Minkowski sum, like that in Figure 3.5, in which one of the summands fails to be convex. Note that the upper boundary of the trajectory $\mathcal{A}$ is concave, since forward motion along that boundary, from right to left, causes the whisker to rotate clockwise. It is this clockwise rotation of Alice's whisker that forces Bob to retreat along the upper, circular boundary of the brush, leading to the retrograde motion and the swallowtail.

3.5 From Minkowski sums to convolutions

A Minkowski sum is simply a set — a subset of the plane, in the case of Figure 3.5. So the boundary of the Minkowski sum $\mathcal{A} \oplus \mathcal{B}$ does not include the swallowtail. Instead, to construct the upper boundary of the set $\mathcal{A} \oplus \mathcal{B}$, we must compute the point of self-intersection of the upper boundary and clip off the swallowtail rooted there, replacing it with a sharp turn.

But the points inside the swallowtail do have a special property that could justify our viewing them as lying inside the sum twice. As we translate the disk-shaped brush $\mathcal{B}$ along the parabolic trajectory $\mathcal{A}$, the points in that swallowtail are precisely the points that get two coats of paint. Given any such point $\mathbf{p}$, the brush moves over $\mathbf{p}$ and then off of $\mathbf{p}$ twice, once while the brush is heading down and again while it is heading back up. This could justify retaining the swallowtail as part of the boundary — in which case the object that it bounds must be a painting, rather than a region. The painting that results when we include the entire offset curve as part of the boundary, retaining the swallowtail, is called the *convolution* of the regions $\mathcal{A}$ and $\mathcal{B}$, written $\mathcal{A} * \mathcal{B}$. Note that the values of the convolution painting $\mathcal{A} * \mathcal{B}$ in Figure 3.5 count coats of paint, just as do the values of the alpha-shaped painting in Figure 3.1.

When winding numbers greater than 1 arise, so that the convolution differs from the Minkowski sum, it is the convolution that is more closely related to fiber products. Note that, as we move along the boundary of the convolution $\mathcal{A} * \mathcal{B}$, the whisker vector changes continuously. On the boundary of the Minkowski sum $\mathcal{A} \oplus \mathcal{B}$, on the other hand, at the throat of the swallowtail, the whisker vector jumps discontinuously. So we shall switch our attention, in what follows, from Minkowski sums to convolutions.

There is another way to count the coats of paint that a point $\mathbf{p}$ receives, and this helps to explain why the name "convolution" is appropriate. Instead of translating the brush $\mathcal{B}$ along the trajectory $\mathcal{A}$ and watching what happens at the point $\mathbf{p}$ during this process, we center a copy of the brush $\mathcal{B}$ at $\mathbf{p}$, intersect that copy with the trajectory $\mathcal{A}$, and count the number of closed segments in the resulting intersection. Note that each such closed segment corresponds to a time interval during which one coat of paint gets applied.

That recipe actually exploits a symmetry of this particular brush; to get the right answer for an arbitrary brush shape $\mathcal{B}$, it turns out that we must invert the brush $\mathcal{B}$ through the origin before translating it to $\mathbf{p}$ and then intersecting it with the trajectory $\mathcal{A}$. (Recall that inversion through the origin, in the plane, is the same as rotating around the origin by 180 degrees; but inversion through the origin in a space of odd dimension is a rotary reflection, rather than a rotation.)

In symbols, this new way of counting coats of paint can be written

$$(\mathcal{A} * \mathcal{B})(\mathbf{p}) = \int \mathcal{A}(\mathbf{q})\, \mathcal{B}(\mathbf{p}-\mathbf{q})\, d\mathbf{q}. \tag{3-3}$$

The painting $\mathcal{B}(\mathbf{p} - \,\cdot\,)$ is the result of inverting $\mathcal{B}$ through the origin and then translating the result to $\mathbf{p}$. We multiply that painting pointwise by the painting $\mathcal{A}(\,\cdot\,)$, this multiplication generalizing the notion of intersection in the description above. We then reduce the resulting product painting to a single integer by using an appropriate reduction map, written here as integration.

Formula 3-3 helps to explain why the name "convolution" is appropriate. If $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ were, say, two sequences of real numbers, their convolution $a * b$ would be defined by the analogous formula

$$(a * b)_n = \sum_i a_i b_{n-i}\,.$$

In both cases, we invert one factor through the origin, translate it to the point of interest, multiply it by the other factor pointwise, and then reduce the result to a single value using some linear operator. Thus, it makes good sense to view the variant of Minkowski sum that retains the swallowtail as a kind of convolution.

It doesn't matter, by the way, which operand gets inverted and translated. It would work equally well to invert the trajectory through the origin, translate it to the point $\mathbf{p}$, and then multiply that by the original brush:

$$(\mathcal{A} * \mathcal{B})(\mathbf{p}) = \int \mathcal{A}(\mathbf{p}-\mathbf{q})\, \mathcal{B}(\mathbf{q})\, d\mathbf{q}.$$

For more details about the reduction operation for paintings, the operator that we are writing as integration, see Schapira [15]. In brief, what that operator counts is the topological degree of the whisker map. Thus, given a painting P, say in the plane, we compute the integer $\int P$ by walking around the entire boundary of P once and counting the net number of counterclockwise full turns that that whisker vector makes during this process. If a painting P is the characteristic function of a

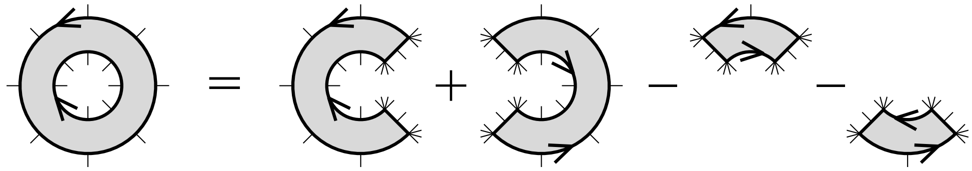


Figure 3.6: Decomposing an annulus into four compact, contractible chunks

set that is both compact and contractible, then the whisker vector will make a net rotation of precisely one full turn, as we move around the boundary of that set; so the integral will be $\int P = 1$. When we apply Formula 3-3 to count coats of paint, the paintings $\mathcal{A}(\mathbf{q})\,\mathcal{B}(\mathbf{p}-\mathbf{q})\,d\mathbf{q}$ that we integrate are the characteristic functions of unions of closed segments of the trajectory $\mathcal{A}$. Since each such closed segment is compact and contractible, the resulting integral is simply the number of segments, which is the number of coats of paint.

Exercise 3-4 Contractibility is important. Let the painting P be the characteristic function of a closed annulus in the plane; what is $\int P$?

Answer: The whisker vector rotates through one positive full turn as we trace the outer boundary of the annulus; but it rotates through one negative full turn as we trace the inner boundary. So we have $\int P = 0$. To see this another way, Figure 3.6 expresses the annulus P as an integral linear combination of four paintings, each of which is the characteristic function of a compact and contractible set. Each summand has integral 1, so the integral of P is $1 + 1 - 1 - 1 = 0$.

3.6 Open versus closed boundaries

Something new happens along the bottom edge of the swallowtail in Figure 3.5: The orientation of that curved edge as encoded by the arrowheads differs from its orientation as encoded by the whiskers. In particular, the whiskers along that edge point toward the left-hand barb of the arrowhead. Everywhere else in all of our examples so far, the whiskers have pointed toward the right-hand barb.

Did we screw up somehow? Is the arrowhead along the bottom edge of the swallowtail correct as drawn? Recall that we are using arrowheads to compute winding numbers. The arrowheads on the other portions of the boundary already imply that the winding number inside the swallowtail must be 2. As a result, the arrowheads along the bottom edge of the swallowtail must point from left to right, to orient that segment of boundary with the higher winding number to its left.

What about the whiskers along the bottom edge of the swallowtail? Are they correct as drawn, pointing upwards? That bottom edge is generated by sums of the form $\mathbf{a} + \mathbf{b}$, where $\mathbf{a}$ is advancing slowly leftward along the middle of the upper boundary of the trajectory $\mathcal{A}$, while $\mathbf{b}$ is retreating quickly rightward around the upper boundary of the brush $\mathcal{B}$. The whiskers at the points $\mathbf{a}$ and $\mathbf{b}$ both point up, so it would certainly be simplest if the whiskers generated along the bottom edge of the swallowtail pointed up as well. Indeed, we can argue that they must point up by comparing the various offsets of the parabola shown in Figure 3.4. Letting $\mathcal{B}_r$ denote a circular brush of radius r, we certainly want to have

$$\mathcal{A} * \mathcal{B}_{6/5} = \mathcal{A} * (\mathcal{B}_{4/5} * \mathcal{B}_{2/5}) = (\mathcal{A} * \mathcal{B}_{4/5}) * \mathcal{B}_{2/5};$$

that is, thickening the trajectory first by $4/5$ and then by an additional $2/5$ should give the same result as a single thickening by the overall distance of $6/5$. For that to hold, the whiskers along the bottom edge of a swallowtail must point up, so that those points are available to be added to points along the top of a subsequent brush, thereby filling in the bottom edge of a larger, higher-up swallowtail.

So we have made no mistake: The orientations on the boundary of a painting that come from the arrowheads and from the whiskers, while they often agree, can sometimes disagree. It follows that we must allow for boundary segments of two different types in our paintings: segments along which the angle between the arrowhead and the whisker is -90 and segments along which that angle is $+90$. The sign of that angle constitutes, in some sense, one additional bit of information that is associated with each segment of the boundary.

That additional bit can be used to encode useful information; and that insight seems worth discussing here briefly, even though it is peripheral to our current purposes. In particular, we can use that bit to distinguish between segments of boundary that are open versus closed. That is, the bit can tell us whether the points that lie precisely on a segment of boundary have the larger or the smaller of the two adjacent winding numbers. Under the simplest encoding scheme, the whiskers point toward the right-hand barb along closed boundaries, but toward the left-hand barb along open boundaries. Most of our paintings have had most of their boundaries closed. For example, the disk-shaped brush $\mathcal{B}$ in Figure 3.5 has winding number 1 along its circular boundary, as well as inside that circle. This convention corresponds to the following rule for computing winding numbers along a boundary:

> To compute the winding number at a point $\mathbf{p}$ that lies precisely on the boundary, perturb the boundary everywhere slightly by pulling on all of its whiskers, and then compute the winding number of that perturbed boundary around $\mathbf{p}$.

For the disk-shaped brush $\mathcal{B}$ in Figure 3.5, the whiskers point outward. Pulling on those whiskers moves the boundary slightly out, increasing its radius; so the points lying precisely on the original boundary get the same winding number as the points inside it.

Under this convention, Figure 3.1(c) shows a closed alpha and Figure 3.2 shows two closed rods whose Minkowski sum is a closed, rounded rectangle. The first segment of open boundary that arises in our figures is the bottom edge of the swallowtail in Figure 3.5. Since the whiskers along that curved edge point up, pulling on the whiskers lifts that segment of boundary slightly; so the points along that curve get winding number 1, rather than 2. Note that this is consistent with counting coats of paint. A point **p** along that bottom edge receives just one coat of paint as the disk-shaped brush $\mathcal{B}$ slides along the parabolic trajectory $\mathcal{A}$. The brush first covers **p**, then almost uncovers **p** — moving so as to bring **p** under some point of the bounding circle of $\mathcal{B}$ — but then covers **p** again, before finally uncovering **p** definitively. So, overall, just one coat of paint.

This talk of open versus closed boundaries, while interesting, is peripheral to our current purpose. The key point for us is that left-barb segments of boundary do arise, as well as right-barb segments. Hence, when computing convolutions, we must keep track of the arrowheads on our boundaries as a separate structure, independent from the whiskers. The natural way to keep track of those arrowheads is to treat our boundaries as oriented manifolds. So the boundaries A and B of the regions $\mathcal{A}$ and $\mathcal{B}$ will come to us with preferred orientations. We want the fiber product $A \times_{\mathbb{S}^1} B$, which is the boundary of the convolution $\mathcal{A} * \mathcal{B}$, to inherit a preferred orientation as well. Thus, convolutions are one application where it is important that the transverse fiber products of smooth, oriented manifolds should themselves be oriented in some natural way. That is the topic of this monograph.

Note that the fiber products that arise from convolving paintings are always equidimensional. If we are dealing with paintings $\mathcal{A}$ and $\mathcal{B}$ in d-space, then their boundary manifolds A and B have dimension $d-1$, while the base manifold is simply the sphere $\mathbb{S}^{d-1}$ of all possible whisker vectors. Thus, readers who are motivated purely by their desire to implement convolution can focus on our equidimensional results, ignoring the large chunk of this monograph that tackles the subtleties of the any-dimensional case.

By the way, we want convolution to be both associative and commutative: to satisfy both $\mathcal{A} * (\mathcal{B} * \mathcal{C}) = (\mathcal{A} * \mathcal{B}) * \mathcal{C}$ and $\mathcal{A} * \mathcal{B} = \mathcal{B} * \mathcal{A}$, for all paintings $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$. So whatever rule we come up with for orienting fiber products should also be both associative and commutative, at least in the equidimensional case. Once we leave the equidimensional case, it turns out that commutativity becomes hopeless; but our final orientation rule will be associative, without any restrictions on the dimensions — indeed, associative even for mixed fiber products.

The convolution operation also has an identity element: the painting $\mathcal{I}$ that interprets the origin of d-space as a d-ball of radius zero. The boundary I of this painting $\mathcal{I}$ is a $(d-1)$-sphere whose whisker map $w_I : I \to \mathbb{S}^{d-1}$ is the identity. To arrange that $\mathcal{I} * \mathcal{A} = \mathcal{A} * \mathcal{I} = \mathcal{A}$ for any painting $\mathcal{A}$, we want the boundaries to satisfy $I \times_{\mathbb{S}^{d-1}} A = A \times_{\mathbb{S}^{d-1}} I = A$. Our final orientation rule will be required to achieve this by the Left and Right Identity Axioms 9.1.5 and 9.1.6.

3.7 Boundaries with more than one component

Way back in Chapter 1, while we were studying Alice and Bob, we discovered that the fiber products of connected manifolds need not be connected. Figure 3.7 shows an example of this phenomenon: two check marks, $\mathcal{L}$ and $\mathcal{R}$, each a non-

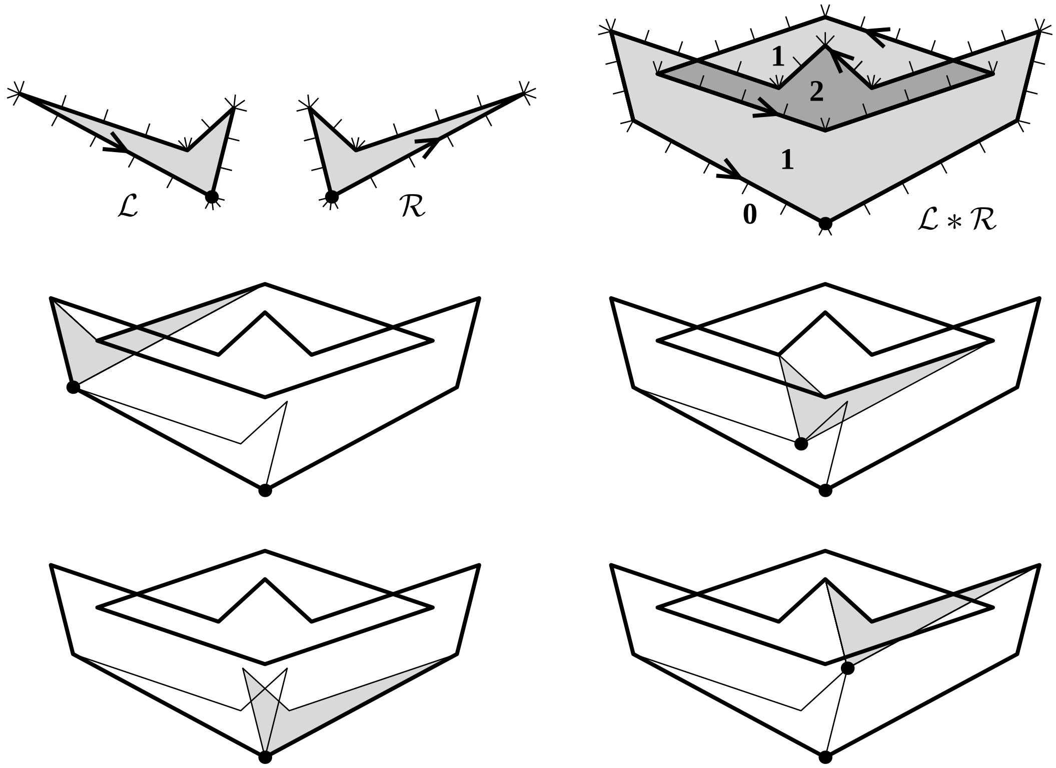


Figure 3.7: The convolution of two non-convex quadrilaterals

convex quadrilateral, whose convolution looks something like a stealth airplane. To help clarify what is going on in this example, the convolution is shown four more times underneath, with the origin of the right-handed check $\mathcal{R}$ aligned, in turn, with each of the four vertices of the left-handed check $\mathcal{L}$.

Suppose that we start off with Alice at the origin of $\mathcal{L}$ and Bob at the origin of $\mathcal{R}$, at the whiskers that point straight down. As they turn and move, advancing or retreating as specified by the Partner's-Slope Rule, the sums of their locations trace out the wings of the airplane: a non-convex octagon. That octagon forms part of the boundary of the convolution $\mathcal{L} * \mathcal{R}$, but not all of it. We can also start out with Alice at the leftmost vertex of $\mathcal{L}$ and Bob at the rightmost vertex of $\mathcal{R}$, at the whiskers that point straight up. As they turn and move starting from there, the sums of their locations trace out the tail of the airplane: a rhomb. The boundary of the convolution $\mathcal{L} * \mathcal{R}$ includes both the octagon and the rhomb.

The region where the rhomb overlaps the octagon has winding number 2 in the convolution $\mathcal{L} * \mathcal{R}$. There is a good geometric reason for this; but, since neither $\mathcal{L}$ nor $\mathcal{R}$ is 1-dimensional, we can't treat either of them as a trajectory, so we can't justify that winding number by counting coats of paint. Instead, we have to apply Formula 3-3 directly:

$$(\mathcal{L} * \mathcal{R})(\mathbf{p}) = \int \mathcal{L}(\mathbf{q})\, \mathcal{R}(\mathbf{p} - \mathbf{q})\, d\mathbf{q}.$$

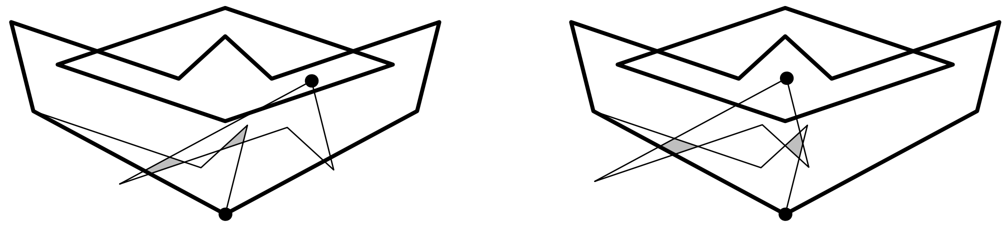


Figure 3.8: Both upper points **p** satisfy $(\mathcal{L} * \mathcal{R})(\mathbf{p}) = 2$.

Figure 3.8 shows what happens when we use this formula to compute the winding number of the convolution $\mathcal{L} * \mathcal{R}$ around two different points **p**, each of which lies in the region of winding number 2, the region where the rhomb overlaps the octagon. In each case, we take the right-handed check $\mathcal{R}$, we invert it through the origin, and we translate the result to bring the origin to the point **p**. The intersection of the inverted and translated $\mathcal{R}$ with the original, left-handed check $\mathcal{L}$ then has two connected components, shown shaded in each half of Figure 3.8. Both of those components are compact and contractible, so the integral is 2.

Note that the two lower edges of the rhomb are segments of open boundary, segments where the whiskers point toward the left-hand barb of the arrowhead, so the points on the boundary have the lower of the two adjacent winding numbers. This openness also makes good geometric sense. If we invert the right-handed check $\mathcal{R}$ and translate it to some point **p** along one of those bottom rhomb edges, the intersection of the result with the original, left-handed check $\mathcal{L}$ will have just one connected component. That intersection will have a cut-point — a single point whose removal would disconnect what remained; but, with the cut-point in place, the intersection is a single component, both compact and contractible.

Chapter 4

Fiber products

Enough, already, of introductions; it is time to start studying fiber products in earnest. The fiber product is most easily thought of as a certain subset of the direct product. So we begin by reviewing the direct products of sets.

4.1 Direct products of sets

When taking direct products of sets, we can keep track of which factor is which in two different ways: by the position of that factor or by an associated index value. Let's first consider the positional approach.

4.1.1 The positional approach

Given two sets A and B, their direct product $A \times B$ is defined, positionally, as the set of all ordered pairs $(\mathbf{a}, \mathbf{b})$, with $\mathbf{a}$ in A and $\mathbf{b}$ in B. More generally, given n sets A_1 through A_n, the n-ary direct product $A_1 \times \cdots \times A_n$ is the set of all ordered n-tuples $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, with $\mathbf{a}_i$ in A_i, for i from 1 to n.

This positional direct-product operator is almost associative, but not quite. From three sets A, B, and C, we can form three different direct products: the left-associated, nested binary product $(A \times B) \times C$, the right-associated analog $A \times (B \times C)$, and the ternary product $A \times B \times C$. Those three sets are distinct; but the reparenthesizing maps $((\mathbf{a}, \mathbf{b}), \mathbf{c}) \leftrightarrow (\mathbf{a}, (\mathbf{b}, \mathbf{c})) \leftrightarrow (\mathbf{a}, \mathbf{b}, \mathbf{c})$ give us obvious one-to-one correspondences between them. Hence, it is generally safe to ignore the distinctions between those three sets and to treat the direct-product operator as if it were associative.

What about commutativity? Unless the sets A and B are equal, the set of ordered pairs $A \times B$ is different from the set $B \times A$, which suggests that the direct-product operator is not commutative. But we have just agreed to ignore the distinction between the sets $(A \times B) \times C$ and $A \times (B \times C)$, because of the natural one-to-one correspondence between them. Should we also, perhaps, ignore the distinction between $A \times B$ and $B \times A$, exploiting the swapping map

$(\mathbf{a}, \mathbf{b}) \leftrightarrow (\mathbf{b}, \mathbf{a})$ as our one-to-one correspondence? But doing that would lead to confusion when the sets A and B intersect. If $\mathbf{p}$ and $\mathbf{q}$ are distinct elements of the intersection $A \cap B$, we must distinguish between the pairs $(\mathbf{p}, \mathbf{q})$ and $(\mathbf{q}, \mathbf{p})$; indeed, that is why we adopted ordered pairs in the first place.

4.1.2 The indexed approach

Commutativity causes confusion in the positional approach because position is a poor way to keep track of which component is which; indices are better. We start with sets A_i, for each i in some index set I. We then define the direct product $\prod_{i \in I} A_i$ to be the set of all functions $\mathbf{a}$ with domain I and whose value $\mathbf{a}_i$ lies in A_i, for all i in I. Note that an ordered pair is essentially the same thing as a function on the domain $\{1, 2\}$; so the indexed approach is not that different from the positional approach. The key advantage of the indexed approach is that we use the same index, say i in I, to name both the factor set A_i and the corresponding component $\mathbf{a}_i$ of an element $\mathbf{a}$ of the direct product.

In the indexed approach, how close is the direct-product operator to being associative? Suppose that we have some partition $I = \bigcup_{k \in K} I_k$ of the index set I into disjoint subsets I_k, for k in another index set K. The direct-product operator is almost associative in the sense that the nested product $\prod_{k \in K} \left(\prod_{i \in I_k} A_i \right)$ is essentially the same as the overall product $\prod_{i \in I} A_i$. An element of the overall product is a function on I whose value at i lies in A_i, while an element of the nested product is a function on K whose value at k is a function on I_k whose value at i lies in A_i. Those two concepts aren't identical, but we can freely convert any element of either product into a corresponding element of the other product. So it is generally safe to ignore the distinction between the two.

As for commutativity, the indexed approach makes it clear that the direct-product operator is commutative, in the sense that it doesn't matter how the index set I is ordered. Indeed, there is no reason for the index set I to be ordered at all.

Another advantage of the indexed approach is that it extends to handle index sets I that are infinite, even uncountably infinite. But that particular advantage isn't relevant for this monograph, since all of the index sets in our direct products — and in our fiber products — are going to be finite. By restricting ourselves to finite index sets, we avoid various set-theoretic subtleties; note that the Axiom of Choice is equivalent to the claim that every direct product of nonempty sets is nonempty, no matter how large the index set of that direct product might be.

4.2 Fiber products of sets

A simple and concrete way to think of the fiber product is as a certain subset of the direct product.

Let S be a fixed set, which we call the *base set*, and let I be any nonempty index set. (Eventually, we want to allow the index set to be empty. But that leads to complications, as we discuss in Section 4.3. For now, let's require I to be nonempty.) For each i in I, let A_i be a set and let $f_i\colon A_i \to S$ be a map; we call A_i the i^{th} *factor set* and f_i the i^{th} *factor map*. Taking the fiber product involves constructing a certain set P and a certain map $h\colon P \to S$.

Let $D := \prod_{i \in I} A_i$ be the direct product of the factor sets, which, under the indexed approach to the direct product, is a set of functions with domain I. The fiber product is the subset P of D consisting of those functions $\mathbf{a}$ that satisfy $f_i(\mathbf{a}_i) = f_j(\mathbf{a}_j)$, for all i and j in I. That is, the fiber product is that subset of the direct product on which all of the factor maps agree. We call this set P the *fiber product* of the sets (A_i) *over* S, and we write

$$P = \prod_{i \in I}{}_S\, [\, f_i]\, A_i.$$

When the factor maps (f_i) are clear from the context, we shall abbreviate this formula by eliding them, writing simply

$$P = \prod_{i \in I}{}_S\, A_i.$$

The fiber-product map $h\colon P \to S$ is then defined by setting $h(\mathbf{a}) := f_i(\mathbf{a}_i)$, for each $\mathbf{a}$ in P and for some i in I. Which i in I we choose doesn't matter, since all of the factor maps agree on P. Furthermore, there is some i to choose, because we are requiring the index set I to be nonempty. We call h the *fiber product* of the maps (f_i), and we write

$$h = \prod_{i \in I}{}_S\, f_i.$$

Is this fiber-product operator associative? The answer is essentially yes, as for the direct product. But we must be careful to avoid having any of the index sets that arise be empty. Suppose that the nonempty index set I is partitioned into disjoint, nonempty subsets I_k, so that we have $I = \bigcup_{k \in K} I_k$; note that K will automatically be nonempty, since I is. We compare the nested fiber product

$$\prod_{k \in K}{}_S\, [h_k] \Big(\prod_{i \in I_k}{}_S\, [\, f_i]\, A_i \Big)$$

to the overall fiber product

$$\prod_{i \in I}{}_S\, [\, f_i]\, A_i,$$

where the factor map h_k in the nested case is the fiber-product map associated with the inner fiber product:

$$h_k := \prod_{i \in I_k}{}_S\, f_i.$$

In the nested case, the relations $f_i(\mathbf{a}_i) = f_j(\mathbf{a}_j)$ with i and j in the same subset I_k are enforced when the inner fiber product

$$\prod_{i \in I_k}{}_S\, [\, f_i\,]\, A_i$$

is formed, while the relations with i and j in different subsets are enforced as part of forming the outer fiber product. In the overall case, all of the relations are enforced simultaneously. But the two end results are the same.

The fiber-product operator is also commutative, in the same sense as the direct-product operator: There is no need for the index set I to be ordered.

4.3 The nullary fiber product

One unfortunate aspect of defining the fiber product to be a subset of the direct product, as we did in Section 4.2, is that we get the wrong answer in the *nullary case*, the case in which the index set I is empty. In this section, we discuss an alternative definition of the fiber product, one that gets the correct answer in the nullary case without treating that case specially. But this alternative definition is more complicated. In the end, we choose to fix up the nullary case simply by treating the case $I = \emptyset$ as special.

Note that the nullary fiber product over S is unique. When the index set I is empty, there are no factor sets A_i and no factor maps $f_i\colon A_i \to S$. So the nullary fiber product over S is some definite set P_0 and some definite map $h_0\colon P_0 \to S$.

What would happen if we tried to apply our original definition in the nullary case? We would first construct the nullary direct product D_0. That set is a singleton, its single element being the unique function whose domain is empty. Since there are no constraints $f_i(\mathbf{a}_i) = f_j(\mathbf{a}_j)$ to enforce, we would take the nullary fiber product P_0 to be all of D_0. We would then get stuck trying to define the nullary fiber-product map h_0, there being no obvious way to choose one element in S to which h_0 should map the single element of P_0.

The correct answer, it turns out, is to take the nullary fiber product P_0 over S to be simply $P_0 = S$, while the associated map $h_0\colon P_0 \to S$ is the identity map $1\colon S \to S$. That answer is correct because the set S, mapped to S via the identity map, acts as an identity element for the fiber-product operation. To see this, let I be a nonempty index set and suppose that e is not an element of I. Letting A_e be another name for S and letting $f_e\colon A_e \to S$ be another name for the identity map $1\colon S \to S$, we claim that the augmented fiber product

$$\prod_{i \in I \cup \{e\}}{}_S\, [\, f_i\,]\, A_i$$

is essentially the same as the plain fiber product

$$\prod_{i \in I}{}_S\, [\, f_i\,]\, A_i.$$

Of course, the two sets are not identical. But, given any element of either set, we can produce the corresponding element of the other in a natural way. Given a point $\mathbf{a}'$ in the augmented fiber product, we construct the corresponding plain point $\mathbf{a}$ by setting $\mathbf{a}_i := \mathbf{a}'_i$ for all i in I, simply ignoring the value $\mathbf{a}'_e$. Given a plain point $\mathbf{a}$, we construct the corresponding augmented point by setting $\mathbf{a}'_i := \mathbf{a}_i$ for all i in I and setting $\mathbf{a}'_e := f_i(\mathbf{a}_i)$ for some i in I, where there is some i to choose and which i we choose doesn't matter.

That argument shows why $P_0 = S$ is the proper answer for the nullary fiber product over S. It also indicates how our definition of the fiber product could be altered, if we wished, in order to get the correct answer in the nullary case. We could simply augment all of our fiber products with one additional factor of S. That augmentation would convert an n-ary product into a product of arity $n+1$, after which nullary products would no longer arise.

Here is how that augmentation would work in detail. Let the base set S, the index set I, the factor sets (A_i), the factor maps (f_i), and the direct product $D := \prod_{i \in I} A_i$ be as before. Consider the augmented direct product $\check{D} := S \times D$. We would define the fiber product to be the subset $\check{P}$ of $\check{D}$ consisting of those pairs $(\mathbf{s}, \mathbf{a})$ in which $f_i(\mathbf{a}_i) = \mathbf{s}$, for all i in I. It follows from this that $f_i(\mathbf{a}_i) = f_j(\mathbf{a}_j)$, for all i and j in I. The fiber-product map $\check{h}\colon \check{P} \to S$ is then simply the projection $\check{h}(\mathbf{s}, \mathbf{a}) := \mathbf{s}$. When I is nonempty, the plain and augmented fiber products P and $\check{P}$ are naturally isomorphic. When I is empty, however, the augmented construction shines by getting the correct answers $\check{P} = S$ and $\check{h} = 1_{S \to S}$.

The augmented definition gets the nullary case correct, but it prevents us from thinking of the fiber product as a subset of the direct product. So we shall stick with the plain definition of Section 4.2, fixing up the nullary case by treating it specially. From now on, therefore, we exclude the nullary case unless it is particularly mentioned.

4.4 An aside about category theory

In category theory, direct products and fiber products are almost the same notion: Fiber products over S in a given category $\mathcal{C}$ are simply direct products in the new category $\mathcal{C}_S$ whose objects are $\mathcal{C}$-morphisms from some $\mathcal{C}$-object to S. We review those ideas here briefly, since the perspective of category theory can be enlightening [11]. But some readers might prefer to skip this section.

We first review how direct products are defined in a category $\mathcal{C}$. For each i in some index set I, let A_i be an object of $\mathcal{C}$. We consider structures of the form $(P, (e_i))$, where P is an object of $\mathcal{C}$ and $e_i\colon P \to A_i$ is a morphism of $\mathcal{C}$, for each i in I. Such a structure $(P, (e_i))$ is called a *direct product* of the objects (A_i) when the following universal mapping property holds: Given any other such structure $(P', (e'_i))$, there exist a unique morphism $u\colon P' \to P$ that satisfies $e_i \circ u = e'_i$, for all i in I. Direct products may not exist. When they do exist, however, a straightforward argument shows that they are unique, up to a unique isomorphism;

so it is generally safe to talk about them as if they were uniquely defined.

For example, consider the category whose objects are sets and whose morphisms are maps between sets. In that category, direct products always exist. We could prove this by showing that the concrete direct product defined in Section 4.1 satisfies the universal mapping property and hence qualifies as an abstract direct product — that is, as a direct product in the sense of category theory.

Associativity works out more neatly in this abstract approach. Recall that our concrete direct-product operator on sets is only almost associative: The two sets $(A \times B) \times C$ and $A \times (B \times C)$ are different, although we generally ignore the distinction between them. In category theory, direct products are defined from the start only up to a unique isomorphism. So the extra mechanism needed to ignore the distinction between $(A \times B) \times C$ and $A \times (B \times C)$ comes built in.

What happens when the index set I is empty? The resulting nullary direct product exists just when the category $\mathcal{C}$ has a *terminal* or *universally attracting* object — an object P such that, for all objects P' in $\mathcal{C}$, there exists a unique morphism $u\colon P' \to P$. If a category $\mathcal{C}$ has terminal objects, any two of them are uniquely isomorphic, and the nullary direct product has that common structure. In the category of sets and set-maps, a singleton set is a terminal object. So a nullary direct product of sets is a singleton set.

Fiber products are simply direct products in a different category. Let S be any fixed object of the category $\mathcal{C}$; we call S the *base object*. We build a new category $\mathcal{C}_S$ whose objects are $\mathcal{C}$-morphisms from some $\mathcal{C}$-object to S. If $f\colon A \to S$ and $g\colon B \to S$ are two objects of the new category $\mathcal{C}_S$, a morphism from f to g in $\mathcal{C}_S$ is simply a morphism $h\colon A \to B$ in $\mathcal{C}$ that satisfies $g \circ h = f$. A fiber product over S in the category $\mathcal{C}$ is then just a direct product in the category $\mathcal{C}_S$.

This perspective provides additional evidence about the correct way to define the nullary fiber product over S. That nullary fiber product should be the terminal object of the category $\mathcal{C}_S$. It is easy to check that the identity morphism $1\colon S \to S$ of the category $\mathcal{C}$ is the terminal object of the category $\mathcal{C}_S$.

By the way, binary fiber products arise frequently in category theory, where they are also called *pullbacks*.

4.5 Specializing the indices to be 1 through n

In this monograph, we are restricting the index sets in our direct products and fiber products to be finite. Our goal is to orient the fiber products of oriented manifolds, and orientation is a fundamentally finite-dimensional concept. Hence, the only products that we are going to need are those of finite arity — those in which the index set I is finite.

Given that our index sets are always finite, it is convenient to take, as our standard index set for an n-ary product, the particular set $\{1, \ldots, n\}$, consisting of the first n positive integers. This has several advantages.

One advantage is notational. When we have $I = \{1, \ldots, n\}$, we can write the direct product $\prod_{i\in I} A_i$ in a simpler way, as $A_1 \times \cdots \times A_n$. Similarly, we can write the fiber-product

$$\text{set} \qquad \prod_{i\in I}{}_S [f_i]\, A_i = \prod_{i\in I}{}_S A_i \qquad \text{and map} \qquad \prod_{i\in I}{}_S f_i$$

as $A_1[f_1] \times_S \cdots \times_S [f_n]A_n = A_1 \times_S \cdots \times_S A_n$ and $f_1 \times_S \cdots \times_S f_n$.

For fiber products, taking $I = \{1, \ldots, n\}$ as our standard index set has an additional advantage: It makes it easy to select a nonredundant set of constraints. Recall that the fiber product $A_1 \times_S \cdots \times_S A_n$ is that subset P of the direct product $D := A_1 \times \cdots \times A_n$ consisting of elements $\mathbf{a}$ that have $f_i(\mathbf{a}_i) = f_j(\mathbf{a}_j)$, for all i and j in I. Once n exceeds 2, these constraints are redundant. But when $I = \{1, \ldots, n\}$, we can focus on the adjacent constraints $f_i(\mathbf{a}_i) = f_{i+1}(\mathbf{a}_{i+1})$, for i from 1 to $n-1$, which are nonredundant.

For these reasons, we henceforth take $I = \{1, \ldots, n\}$ as our standard index set for an n-ary product. (As discussed in Section 4.3, we are allowing the nullary case $I = \emptyset$ only when that case is being explicitly discussed. So we typically require that $n \geq 1$.)

One disadvantage of taking $I = \{1, \ldots, n\}$ is that the set $\{1, \ldots, n\}$ has too much structure: In addition to being a set of cardinality n, it is also ordered. Worse yet, we have already exploited that order. We exploited it when we adopted the notations $A_1 \times \cdots \times A_n$ and $A_1 \times_S \cdots \times_S A_n$ as ways to write down the direct and fiber products — notations in which the factor sets are ordered from left to right. We exploited it in another way when, for an n-ary fiber product, we selected a nonredundant set of constraints by focusing on the adjacent pairs of factor maps.

Exploiting the total order on the set $I = \{1, \ldots, n\}$ is bad because it raises the possibility that the resulting product operator might not be commutative — that the product might depend, in some important way, on which ordering of the index set I we choose. Indeed, we shall discuss a product operation in Section 5.2, the oriented direct product of oriented linear spaces, in which the ordering on I does matter. Until then, however, our only uses of the ordering on $I = \{1, \ldots, n\}$ will be unimportant. For example, changing the ordering would change which of the constraints $f_i(\mathbf{a}_i) = f_j(\mathbf{a}_j)$ we chose to focus on as the nonredundant ones; but that change wouldn't affect the resulting fiber product.

4.6 Fiber products of linear spaces

Suppose that the base set S and all of the factor sets A_1 through A_n are linear spaces, say over the real numbers. So the direct product $D := A_1 \times \cdots \times A_n$ is also a linear space; we shall denote an element of that product as $(\alpha_1, \ldots, \alpha_n)$, using Greek letters for the elements of the factor spaces, which we think of as vectors. Suppose also that the factor maps $f_i \colon A_i \to S$, for i from 1 to n, are all linear. Since the constraints $f_i(\alpha_i) = f_{i+1}(\alpha_{i+1})$, for i from 1 to $n-1$, are then linear equations, the fiber product $A_1 \times_S \cdots \times_S A_n$ will be a linear subspace of the direct product $A_1 \times \cdots \times A_n$. The fiber-product map $f_1 \times_S \cdots \times_S f_n$ will also be linear. So we can view the fiber product as an operation in the world of linear spaces and linear maps (more precisely, in that category).

There is another way in which we can exploit linearity. When the base space S has a linear structure, we can measure the extent to which the two adjacent factor maps f_i and f_{i+1} fail to agree by forming the difference $f_{i+1}(\alpha_{i+1}) - f_i(\alpha_i)$, that difference being a vector in the base space S — a vector that we hope is zero. If we take such a difference for each i, from 1 to $n-1$, we end up with a map $\Delta\colon A_1 \times \cdots \times A_n \to S^{n-1}$ that we shall call the *difference map*. (Don't confuse S^{n-1}, the Cartesian product of $n-1$ copies of S, with $\mathbb{S}^{n-1}$, the $(n-1)$-dimensional sphere.) That is, we define the difference map Δ by

$$\Delta(\alpha_1, \ldots, \alpha_n) := \big(f_2(\alpha_2) - f_1(\alpha_1),\ \ldots,\ f_n(\alpha_n) - f_{n-1}(\alpha_{n-1})\big).$$

Note that the fiber product subspace $A_1 \times_S \cdots \times_S A_n$ is precisely the kernel of the difference map Δ; we have $A_1 \times_S \cdots \times_S A_n = \operatorname{Ker}(\Delta)$.

4.7 Transversality

While the fiber product of linear spaces is always a linear space, its dimension depends upon the way in which the factor maps (f_i) interact. For example, if all of the factor maps are identically zero, then each constraint $f_i(\alpha_i) = f_{i+1}(\alpha_{i+1})$ is satisfied trivially, so the fiber product coincides with the direct product. We are particularly interested in the opposite extreme: the *transverse case*, the case in which the fiber product has the smallest possible dimension.

It is convenient to restrict ourselves to linear spaces whose dimension is finite. The notion of transversality can be extended to infinite-dimensional spaces, but doing so requires topological concepts, such as Banach spaces and continuous linear maps, that don't arise in the finite-dimensional case. Since our eventual goal is to study orientation and orientation is a purely finite-dimensional concept, we henceforth restrict ourselves to finite-dimensional spaces.

What will the finite dimension of the fiber product $A_1 \times_S \cdots \times_S A_n$ then be? Let's denote $\dim(A_i)$ by a_i and $\dim(S)$ by s; by convention, we use an upper-case italic letter for a linear space, the matching lower-case italic letter for its dimension, and the corresponding Greek letter for an element of it. For the direct product $D = A_1 \times \cdots \times A_n$, we have $\dim(D) = d = a_1 + \cdots + a_n$. We cut out the fiber product P, inside of D, with the constraints $f_i(\alpha_i) = f_{i+1}(\alpha_{i+1})$, for i from 1 to $n-1$. Each of these constraints is an equality between elements of S, and can hence remove at most s degrees of freedom. The fiber product is as small as possible just when those constraints are independent, jointly removing $(n-1)s$ degrees of freedom and leaving $\dim(P) = p := a_1 + \cdots + a_n - (n-1)s$. Transversality is the property of the factor maps (f_i) that makes those constraints be independent.

Definition 4-1 Let n be a positive integer, let S and A_1 through A_n be finite-dimensional linear spaces, and let $f_i\colon A_i \to S$ be a linear map, for i from 1 to n. Let $\Delta\colon A_1 \times \cdots \times A_n \to S^{n-1}$ be the difference map, defined by

$$\Delta(\alpha_1, \ldots, \alpha_n) := \big(f_2(\alpha_2) - f_1(\alpha_1),\ \ldots,\ f_n(\alpha_n) - f_{n-1}(\alpha_{n-1})\big).$$

The maps $(f_1, \ldots, f_n)$ are called *transverse* just when the difference map Δ is surjective. Note that this can happen only when $\dim(A_1 \times \cdots \times A_n) \geq \dim(S^{n-1})$, that is, when $a_1 + \cdots + a_n \geq (n-1)s$.

Consider a fiber product $P := A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ of linear spaces. It is precisely when the factor maps are transverse that the dimension of P is as small as possible — in particular, that $\dim(P) = a_1 + \cdots + a_n - (n-1)s$. To see why, recall that $P = \mathrm{Ker}(\Delta)$. By elementary linear algebra, we have $\dim(P) = \dim(\mathrm{Ker}(\Delta)) = \dim(A_1 \times \cdots \times A_n) - \dim(\mathrm{Im}(\Delta))$. And the map Δ is surjective just when $\dim(\mathrm{Im}(\Delta)) = \dim(S^{n-1}) = (n-1)s$.

By the way, it is often helpful to think of the minimum possible dimension of the fiber product P as being $s + (a_1 - s) + \cdots + (a_n - s)$.

Exercise 4-2 In the context of Definition 4-1, define the *augmented difference map* to be the map $\check{\Delta}\colon S \times A_1 \times \cdots \times A_n \to S^n$ given by

$$\check{\Delta}(\sigma, \alpha_1, \ldots, \alpha_n) := \big(f_1(\alpha_1) - \sigma,\ \ldots,\ f_n(\alpha_n) - \sigma\big).$$

Show that the map $\check{\Delta}$ is surjective just when Δ is surjective, that is, just when the maps $(f_1, \ldots, f_n)$ are transverse. Note that this augmentation adds a new factor space $A_0 := S$ whose factor map $f_0\colon A_0 \to S$ is the identity map on S, just as we saw in Section 4.3 when studying nullary fiber products. Indeed, this exercise makes it clear that the proper way to extend Definition 4-1 to the nullary case $n = 0$ is to make the convention that the empty sequence of maps is transverse.

Exercise 4-3 If the sequence $(f_1, \ldots, f_n)$ of linear maps is transverse, show that any subsequence, taken in any order, is also transverse. (This follows easily from the previous exercise.)

Exercise 4-4 Show that the linear maps $(f_1, \ldots, f_n)$ are always transverse when all but one of them, say all but f_k, are surjective.

Answer: To show $\check{\Delta}$ surjective, it suffices to start with an arbitrary point $(\sigma_1, \ldots, \sigma_n)$ in S^n and to construct a point $(\sigma, \alpha_1, \ldots, \alpha_n)$ in $S \times A_1 \times \cdots \times A_n$ with $\check{\Delta}(\sigma, \alpha_1, \ldots, \alpha_n) = (\sigma_1, \ldots, \sigma_n)$. We start by fixing α_k arbitrarily and then choosing σ to arrange that $f_k(\alpha_k) - \sigma = \sigma_k$. For i different from k, we can then exploit the surjectivity of f_i to find some α_i with $f_i(\alpha_i) - \sigma = \sigma_i$.

Transversality is particularly simple in the binary case. The difference map $\Delta\colon A_1 \times A_2 \to S$ is then given by $\Delta(\alpha_1, \alpha_2) := f_2(\alpha_2) - f_1(\alpha_1)$, so Δ is surjective just when the images of f_1 and f_2, together, span all of S — that is, just when $\mathrm{Im}(f_1) + \mathrm{Im}(f_2) = S$. Note that the maps (f_1, f_2) will be transverse, in particular, when either $\mathrm{Im}(f_1) = S$ or $\mathrm{Im}(f_2) = S$, as we saw in Exercise 4-4.

Exercise 4-5 Let A, B, and S each be the (x, y) plane, let the factor map $f: A \to S$ be the projection $f(x, y) := (x, 0)$ on the x-axis, and let $g: B \to S$ be the projection $g(x, y) := (0, y)$ on the y-axis. Show that the maps f and g are transverse, even though neither of them is surjective.

Answer: The image of f is the x-axis, the image of g is the y-axis, and those two subspaces, together, span the entire plane S.

Exercise 4-6 Generalize the Invertible Factor Laws, 2-5 through 2-7, from the binary case to the n-ary case.

Answer: The Left and Right Invertible Laws generalize to become the *All-but-One Invertible Law,* which says, for k from 1 to n, that

$$\operatorname{sgn}(u_k) = \operatorname{sgn}(f_1) \cdots \operatorname{sgn}(f_{k-1}) \operatorname{sgn}(f_{k+1}) \cdots \operatorname{sgn}(f_n), \tag{4-7}$$

where $u_k: A_1 \times_S \cdots \times_S A_n \to A_k$ is the projection from the fiber product to the k^{th} factor space. That is, if all of the factor maps but the k^{th} are invertible, then the fiber product is transverse, the k^{th} projection map u_k is invertible, and we should orient the fiber product to make that equality hold.

When all n of the factor maps are invertible, we have the less general but more elegant *All Invertible Law,*

$$\operatorname{sgn}(f_1 \times_S \cdots \times_S f_n) = \operatorname{sgn}(f_1) \cdots \operatorname{sgn}(f_n). \tag{4-8}$$

This law, like the Both Invertible Law 2-7 from the binary case, doesn't need to mention any of the projection maps.

4.7.1 The associativity of transversality

Exercise 4-3 showed that the condition of transversality is commutative; it is also associative, in the following sense. Suppose that we have an overall fiber product whose factor maps may or may not be transverse. We can consider computing that same fiber product in two steps: We first combine some of its factors in an inner fiber product and then combine that inner result with the remaining factors in an outer fiber product. Since the fiber-product operator itself is associative, this two-step process will get the same result as the one-step, overall fiber product. We claim, as well, that the overall product is transverse just when both the inner and outer products are transverse.[1] By commutativity, we can assume that the factors that get combined in the inner fiber product are an initial substring of the factors.

Proposition 4-9 *Let n be a positive integer, let S and A_1 through A_n be finite-dimensional linear spaces, let $f_i: A_i \to S$ be a linear map, for i from 1 to n, and let k be an integer with $1 \le k \le n$. Let P denote the inner fiber product, the product $P := A_1 \times_S \cdots \times_S A_k$ of the first k factor spaces, and let $h: P \to S$*

[1]The phrase "inner product" here is simply an abbreviation of "inner fiber product"; we are not talking about the type of inner product that takes two vectors and returns a scalar.

be the associated fiber-product map $h := f_1 \times_S \cdots \times_S f_k$. *The sequence of maps* $(f_1, \ldots, f_n)$ *associated with the overall fiber product* $A_1 \times_S \cdots \times_S A_n$ *is transverse just when*

- *the sequence of maps* $(f_1, \ldots, f_k)$ *associated with the inner fiber product* $P = A_1 \times_S \cdots \times_S A_k$ *is transverse*
- *and the sequence of maps* $(h, f_{k+1}, \ldots, f_n)$ *associated with the outer fiber product* $P \times_S A_{k+1} \times_S \cdots \times_S A_n$ *is also transverse.*

Proof Let $a_k := \dim(A_k)$ for k from 1 to n and let $s := \dim(S)$. The overall fiber product $A_1 \times_S \cdots \times_S A_n$ has its minimum possible dimension, which is $s + (a_1 - s) + \cdots + (a_n - s)$, just when the maps $(f_1, \ldots, f_n)$ are transverse. On the other hand, suppose that we compute that same product in two steps. The inner product has its minimum possible dimension, $p_{\min} := s + (a_1 - s) + \cdots + (a_k - s)$, just when the maps $(f_1, \ldots, f_k)$ are transverse. And the outer product has its minimum possible dimension, $s + (p - s) + (a_{k+1} - s) + \cdots + (a_n - s)$, just when the maps $(h, f_{k+1}, \ldots, f_n)$ are transverse, where p here denotes the dimension of the inner fiber product. Substituting $p_{\min}$ for p in this expression gives the minimum possible overall dimension; so the overall product is transverse just when both the inner and outer products are transverse. $\square$

4.7.2 Transversality viewed geometrically

The algebraic notion of transversality that we have just defined is closely related to the geometric notion of transverse intersections. Let S be a linear space and let A_1 through A_n be subspaces of S. The subspaces A_1 through A_n are said to *intersect transversely* when the codimension of their intersection is the sum of their codimensions:

$$s - \dim(A_1 \cap \cdots \cap A_n) = (s - a_1) + \cdots + (s - a_n). \tag{4-10}$$

For example, the three planes $x = 0$, $y = 0$, and $z = 0$ in (x, y, z)-space intersect transversely; each plane has codimension $3 - 2 = 1$, while their intersection, which is just the origin, has codimension $3 - 0 = 3 = 1 + 1 + 1$.

To relate this concept to fiber products, we view each subspace A_i as a factor space by equipping it with the identity injection $f_i \colon A_i \to S$ as its factor map. A point $(\alpha_1, \ldots, \alpha_n)$ in the direct product $D = A_1 \times \cdots \times A_n$ then belongs to the fiber product P just when $\alpha_1 = \cdots = \alpha_n$, so the fiber product is essentially the intersection. This fiber product is transverse just when

$$\dim(A_1 \times_S \cdots \times_S A_n) = \dim(A_1 \cap \cdots \cap A_n) = s + (a_1 - s) + \cdots + (a_n - s),$$

which, by Equation 4-10, is just when the subspaces intersect transversely.

Note that, by a standard convention, intersecting zero subspaces of S gives us back simply S itself. This is further evidence that we did the right thing when we defined the nullary fiber product over S to be S.

Now is a convenient time to point out that the transversality of an n-tuple of maps $(f_1, \ldots, f_n)$ is a stronger condition than the transversality of all pairs (f_i, f_j). Consider the three planes $x = 0$, $y = 0$, and $x = y$ in (x, y, z)-space. Each plane has codimension 1, and each pair of planes intersect transversely in a line, of codimension 2. But all three planes also intersect in the entire line $x = y = 0$, so the three planes do not intersect transversely.

4.8 Fiber products of smooth manifolds

A smooth manifold looks everywhere locally like a linear space, so it is not hard to extend the theory of fiber products from linear spaces to smooth manifolds.

A *topological d-manifold* is a Hausdorff topological space in which every point has a neighborhood homeomorphic to an open subset of $\mathbb{R}^d$. If $\mathbf{M}$ is a topological d-manifold, a homeomorphism $\varphi\colon U \to \mathbb{R}^d$ from an open set $U \subseteq \mathbf{M}$ to the open set $\varphi(U) \subseteq \mathbb{R}^d$ is called a *local coordinate system* or *chart* on U. A family of charts that cover the entire manifold $\mathbf{M}$ is called an *atlas*.

(A mathematical fine point: We are going along with most authors in requiring that our topological manifolds be Hausdorff spaces. But Lang [9] points out that this assumption plays no role in the bulk of the arguments involving manifolds. Consider the real line with its origin doubled. Viewed as a topological space, this space is T_1, but is not Hausdorff: Each of the two origins has a neighborhood that does not contain the other; but we can't find neighborhoods of the two origins that are disjoint. Lang considers this space to be a smooth 1-manifold, while most authors do not. This monograph supports Lang's position, in the sense that none of our arguments that involve manifolds require the Hausdorff property. Thus, if you so choose, you are free to follow Lang in allowing non-Hausdorff manifolds.)

A *smooth d-manifold* is a topological d-manifold together with an atlas in which, for every pair of charts $\varphi\colon U \to \mathbb{R}^d$ and $\psi\colon V \to \mathbb{R}^d$, the composite homeomorphism $\psi \circ \varphi^{-1}\colon \varphi(U \cap V) \to \psi(U \cap V)$, called the *change of coordinates map* on the intersection of the two charts, is infinitely differentiable. If $\mathbf{M}$ is a smooth d-manifold and $\mathbf{N}$ is a smooth e-manifold, then a continuous map $\mathbf{f}\colon \mathbf{M} \to \mathbf{N}$ is called *smooth* at the point $\mathbf{m}$ in $\mathbf{M}$ when, for some chart $\varphi\colon U \to \mathbb{R}^d$ on a neighborhood U of $\mathbf{m}$ in $\mathbf{M}$ and for some chart $\psi\colon V \to \mathbb{R}^e$ on a neighborhood V of $\mathbf{f}(\mathbf{m})$ in $\mathbf{N}$, the composite map

$$\psi \circ \mathbf{f} \circ \varphi^{-1}\colon \varphi(U \cap \mathbf{f}^{-1}(V)) \to \psi(V)$$

is infinitely differentiable, as a map from an open set in $\mathbb{R}^d$ to $\mathbb{R}^e$. This condition doesn't depend upon which charts φ and ψ we choose. The map $\mathbf{f}\colon \mathbf{M} \to \mathbf{N}$ is *smooth* when $\mathbf{f}$ is smooth at all points $\mathbf{m}$ in $\mathbf{M}$. (We are being lazy in taking "smooth" to mean infinitely differentiable, that is, C^∞. For our purposes, C^1 would probably be smooth enough; but it is convenient not to have to keep track of the number of continuous derivatives.)

So we consider smooth manifolds and smooth maps between them; what about products in this category? Direct products always exist; but fiber products are more subtle. If we don't impose any restrictions, it might happen, for smooth manifolds $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{S}$ and for smooth maps $\mathbf{f}\colon \mathbf{A} \to \mathbf{S}$ and $\mathbf{g}\colon \mathbf{B} \to \mathbf{S}$, that the local dimension of the set-theoretic fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ varies from point to point, as a consequence of the local behaviors of $\mathbf{f}$ and $\mathbf{g}$. Indeed, we saw an example of essentially this phenomenon in Figure 1.6, where the local dimension of a non-transverse fiber product is sometimes 1 and sometimes 2. We are going to rule out this bad phenomenon in the standard way by requiring, at every point on the set-theoretic fiber product, that the linear approximations to the factor maps $\mathbf{f}$ and $\mathbf{g}$ be transverse. This forces the local dimension of the fiber product to be everywhere minimal — and hence to be everywhere the same.

We recall some notation. Let $\mathbf{M}$ and $\mathbf{N}$ be smooth manifolds, let $\mathbf{f}\colon \mathbf{M} \to \mathbf{N}$ be a smooth map, and let $\mathbf{m}$ be a point in $\mathbf{M}$. We denote by $T_{\mathbf{m}}\mathbf{M}$ the tangent space to the manifold $\mathbf{M}$ at the point $\mathbf{m}$. And we denote by $T_{\mathbf{m}}\mathbf{f}\colon T_{\mathbf{m}}\mathbf{M} \to T_{\mathbf{f}(\mathbf{m})}\mathbf{N}$ the differential of $\mathbf{f}$ at $\mathbf{m}$, which is the linear map from the tangent space $T_{\mathbf{m}}\mathbf{M}$ to the tangent space $T_{\mathbf{f}(\mathbf{m})}\mathbf{N}$ that approximates $\mathbf{f}$ to first order near $\mathbf{m}$. The notations $(d\mathbf{f})_{\mathbf{m}}$ and $\mathbf{f}'(\mathbf{m})$ are also used for the differential $T_{\mathbf{m}}\mathbf{f}$.

Definition 4-11 Let n be a positive integer and let $\mathbf{S}$ and $\mathbf{A}_1$ through $\mathbf{A}_n$ be finite-dimensional smooth manifolds. For i from 1 to n, let $\mathbf{f}_i\colon \mathbf{A}_i \to \mathbf{S}$ be a smooth map. The maps $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ are called *transverse* when, for every point $\mathbf{s}$ in $\mathbf{S}$ and for every n-tuple of points $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ with $\mathbf{a}_i$ in $\mathbf{A}_i$ and $\mathbf{f}_i(\mathbf{a}_i) = \mathbf{s}$ for i from 1 to n, the differentials $T_{\mathbf{a}_i}\mathbf{f}_i\colon T_{\mathbf{a}_i}\mathbf{A}_i \to T_{\mathbf{s}}\mathbf{S}$, viewed as a sequence of linear maps $(T_{\mathbf{a}_1}\mathbf{f}_1, \dots, T_{\mathbf{a}_n}\mathbf{f}_n)$, are transverse.

We shall often abbreviate the linear-space fiber product that represents the local behavior of a smooth-manifold fiber product by using italic letters, rather than boldface. That is, we shall abbreviate the tangent space $T_{\mathbf{a}_i}\mathbf{A}_i$ as A_i, the tangent space $T_{\mathbf{s}}\mathbf{S}$ as S, and the linear map $T_{\mathbf{a}_i}\mathbf{f}_i$ as $f_i\colon A_i \to S$, leading to the linear-space fiber product $A_1[f_1] \times_S \cdots \times_S [f_n]A_n$.

When the smooth-manifold fiber product $\mathbf{A}_1[\mathbf{f}_1] \times_{\mathbf{S}} \cdots \times_{\mathbf{S}} [\mathbf{f}_n]\mathbf{A}_n$ is transverse, the linear-space fiber product $A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ that represents the local behavior at any point is always transverse. It follows that the local dimension of the fiber product $\mathbf{A}_1 \times_{\mathbf{S}} \cdots \times_{\mathbf{S}} \mathbf{A}_n$, at every point on it, is the minimum possible, that minimum dimension being $(a_1 + \cdots + a_n) - (n-1)s$. As a result, the fiber product $\mathbf{A}_1 \times_{\mathbf{S}} \cdots \times_{\mathbf{S}} \mathbf{A}_n$ is again a smooth manifold.

Proposition 4-12 *With the same notations as in Definition 4-11, suppose that the smooth factor maps* $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ *are in fact transverse. The set-theoretic fiber product* $\mathbf{P} := \mathbf{A}_1[\mathbf{f}_1] \times_{\mathbf{S}} \cdots \times_{\mathbf{S}} [\mathbf{f}_n]\mathbf{A}_n$ *is then a smooth manifold, of dimension* $(a_1 + \cdots + a_n) - (n-1)s$. *Also, the fiber-product map* $\mathbf{f}_1 \times_{\mathbf{S}} \cdots \times_{\mathbf{S}} \mathbf{f}_n\colon \mathbf{P} \to \mathbf{S}$ *and the projection maps* $\mathbf{u}_i\colon \mathbf{P} \to \mathbf{A}_i$, *for* i *from* 1 *to* n, *are smooth maps.*

Proof This is a standard result, for which we appeal to standard texts, such as Lang [10] — but beware that the proof given in Lang consists of the single word "Obvious." □

Beware also: It can easily happen that the fiber product $\mathbf{A}_1 \times_{\mathbf{S}} \cdots \times_{\mathbf{S}} \mathbf{A}_n$ is the empty set, that is, that there are no points $\mathbf{s}$ in $\mathbf{S}$ and n-tuples of points $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ with $\mathbf{a}_i$ in $\mathbf{A}_i$ and $\mathbf{f}_i(\mathbf{a}_i) = \mathbf{s}$, for i from 1 to n. In this case, the maps $(\mathbf{f}_1, \ldots, \mathbf{f}_n)$ are transverse vacuously. Indeed, when $a_1 + \cdots + a_n < (n-1)s$, the only way that the maps $(\mathbf{f}_1, \ldots, \mathbf{f}_n)$ can be transverse is vacuously, since we know that we must have $a_1 + \cdots + a_n \geq (n-1)s$ in order for any linear-space fiber product with those dimensions to be transverse.

Suppose that the maps $(\mathbf{f}_1, \ldots, \mathbf{f}_n)$ are transverse vacuously. Proposition 4-12 then tells us that the fiber product $\mathbf{A}_1 \times_{\mathbf{S}} \cdots \times_{\mathbf{S}} \mathbf{A}_n = \emptyset$ is a smooth manifold and that it has dimension $d := (a_1 + \cdots + a_n) - (n-1)s$. That's fine when $d \geq 0$; everyone agrees that the empty set is a perfectly valid smooth manifold of any nonnegative dimension d. But what about negative dimensions? To avoid a special case in Proposition 4-12, we make the convention that the empty set is a smooth manifold of dimension d also when $d < 0$. Of course, there are no linear spaces of negative dimension; but an empty manifold has no tangent spaces, so this convention is defensible.

Proposition 4-12 tells us that fiber products do exist in the category of smooth manifolds, as long as we restrict ourselves to the transverse case. Much of the theory then carries over from linear spaces to smooth manifolds. For example, the associativity of the fiber-product operator and the associativity of the condition of transversality both carry over from linear spaces to smooth manifolds.

Geometers distinguish certain sequences of submanifolds $(\mathbf{A}_1, \ldots, \mathbf{A}_n)$ of a smooth manifold $\mathbf{S}$ as *intersecting transversely*. The relationship of this geometric notion to the map-based notion of transversality in Definition 4-11 is the same for manifolds as for linear spaces: We let the factor maps be the identity injections $\mathbf{f}_i \colon \mathbf{A}_i \to \mathbf{S}$. Two smooth surfaces in a smooth 3-fold intersect transversely just when, at each point of intersection, the two tangent planes do not coincide. A curve and a surface intersect transversely in a 3-fold just when, at each point of intersection, the tangent line to the curve does not lie in the tangent plane to the surface. Two curves can intersect transversely in a 3-fold only by being skew — that is, by having no points of intersection.

If $\mathbf{A}$ and $\mathbf{B}$ are two smooth curves that lie skew in a smooth 3-fold $\mathbf{S}$, then the intersection $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ is an example of a transverse fiber product of negative dimension; we have $\dim(\mathbf{A} \times_{\mathbf{S}} \mathbf{B}) = \dim(\mathbf{A}) + \dim(\mathbf{B}) - \dim(\mathbf{S}) = 1 + 1 - 3 = -1$. There is no contradiction here, because the intersection $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ is empty.

Exercise 4-13 While we shan't prove Proposition 4-12 in this monograph, we should know how to construct charts on the fiber-product manifold. For notational simplicity, let's consider a binary fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$, which we assume transverse. Let $\mathbf{p} = (\mathbf{a}, \mathbf{b})$ be a point in the fiber product, so that $\mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b})$

is some point **s** in **S**. There is some neighborhood of the point **a** in the manifold **A** that is homeomorphic to an open subset of $\mathbb{R}^a$. Since we are interested only in what happens near **p**, let's assume for simplicity that the manifold **A** actually coincides with an open subset of the linear space $\mathbb{R}^a$; and let's also translate, if necessary, to get $\mathbf{a} = 0$. Similarly, let's assume that **B** is a neighborhood of the origin $\mathbf{b} = 0$ in $\mathbb{R}^b$ and that **S** is a neighborhood of the origin $\mathbf{s} = 0$ in $\mathbb{R}^s$. Construct a chart on a neighborhood of the point $\mathbf{p} = (0, 0)$ in the fiber product.

Sketch: Given our simplifying assumptions, the fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$ is a submanifold of the linear space $\mathbb{R}^a \times \mathbb{R}^b = \mathbb{R}^{a+b}$. The tangent space to that submanifold at the origin is the linear-space fiber product $T_{(0,0)}(\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}) = \mathbb{R}^a[T_0\mathbf{f}] \times_{\mathbb{R}^s} [T_0\mathbf{g}]\,\mathbb{R}^b$, which has dimension $a + b - s$ by transversality. Let C be some complement of that tangent space in $\mathbb{R}^{a+b}$, and note that $\dim(C) = s$. Convince yourself that the linear-space quotient map from $\mathbb{R}^{a+b}$ to $\mathbb{R}^{a+b}/C$, when restricted to the fiber product $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$, will be a chart on some sufficiently small neighborhood of the origin in the fiber product.

Chapter 5

Orientation

Orientation is the property that, for a curve, distinguishes forward from backward; for a surface, clockwise from counterclockwise; and for a 3-fold, right-handed from left-handed. We first review the concept of orienting linear spaces and then move on to orienting smooth manifolds.

5.1 The sign of an ordered basis

Orienting finite-dimensional linear spaces over the real numbers $\mathbb{R}$ is a standard topic; but we have to be careful here to make sure that every space — including a space of dimension zero — ends up with two possible orientations.

If M is a real linear space whose dimension $m := \dim(M)$ is finite, an *orientation* of M is a rule that assigns a sign to each ordered basis of M in one of the two globally consistent ways. We can encode an orientation of M as a pair $(\mathbf{b}, \beta)$ consisting of some ordered basis $\mathbf{b} := (\mu_1, \ldots, \mu_m)$ for M and a single bit β, say encoded as an element of the set $\{+1, -1\}$. We interpret the pair $(\mathbf{b}, \beta)$ as declaring that the ordered basis $\mathbf{b} = (\mu_1, \ldots, \mu_m)$ has the sign β. Once one basis for M has been assigned a sign, every other basis inherits a unique sign by the consistency constraint: The basis $\mathbf{c}$ has the same sign as $\mathbf{b}$ just when the matrix that transforms from $\mathbf{b}$ to $\mathbf{c}$ — or back again, it doesn't matter — has positive determinant. An *oriented linear space* is a linear space together with a chosen orientation. If M is an oriented linear space, we use $-M$ to denote the same underlying linear space, but equipped with the opposite orientation; so a basis is positive for $-M$ just when it is negative for M.

If $\mathbf{b}$ is an ordered sequence of linearly independent vectors in some linear space, let's write $\langle\mathbf{b}\rangle$ to denote the span of those vectors, oriented so as to make the sequence $\mathbf{b}$ a positive basis for the space $\langle\mathbf{b}\rangle$. So, if $\mathbf{b}$ is a basis for the entire oriented linear space M, we have either $\langle\mathbf{b}\rangle = M$ or $\langle\mathbf{b}\rangle = -M$, according as the basis $\mathbf{b}$ is positive or negative.

A linear space M of dimension zero is special in that it has only one basis: the empty sequence of vectors. Such a space M still has two possible orientations, under our definition, since that unique basis can be either positive or negative, so we have either $M = \langle\rangle$ or $M = -\langle\rangle$. Be warned that people sometimes talk, informally, as if orienting a linear space meant specifying a positive basis. But, if that were the case, then a zero-dimensional space would have only one possible orientation. The existence of two possible orientations turns out to be essential in what follows, since a fiber product $A[f] \times_S [g]B$ can be zero-dimensional even when the input spaces A, B, and S all have positive dimension. Reversing the orientation on any one of A, B, or S (while leaving the maps f and g unchanged) should reverse the induced orientation on the fiber product $A \times_S B$; to make this possible, even a zero-dimensional space must have two possible orientations.

While a zero-dimensional space is just like a space of positive dimension in having two possible orientations, there is still something special about a zero-dimensional space: Its two orientations are intrinsically distinguishable. Any oriented linear space of positive dimension has both positive bases and negative bases. Indeed, if we negate any vector in any positive basis, the result is a negative basis, and vice versa; so there are "just as many" positive bases as negative ones. It doesn't make sense to ask whether such a space is itself positively oriented or negatively oriented, in any intrinsic sense. But let M be an oriented linear space with $\dim(M) = 0$. The space M has exactly one basis: the empty sequence. And that basis is either positive or negative. There are no vectors in that basis to negate, to convert it from one sign to the other. We have either $M = \langle\rangle$ or $M = -\langle\rangle$, and we can tell which of those two cases pertains. Given an oriented linear space M of dimension zero, we'll say that M is *positively oriented* or *negatively oriented* according as $M = \langle\rangle$ or $M = -\langle\rangle$.

Let's adopt the symbol $\Diamond$ as a prettier way to write $\langle\rangle$, the oriented linear span of the empty sequence. So the space $\Diamond$ is positively oriented, while the space $-\Diamond$ is negatively oriented. (More precisely, the symbol "$\Diamond$" denotes the zero-dimensional subspace of the current linear space, where that containing space must be determined from context. This is similar to the way that the symbol "0" is used to denote the zero element of the current linear space, where that space is determined from context. When confusion might arise, we shall write $\Diamond_U$ to denote the zero-dimensional subspace of the linear space U.)

Exercise 5-1 True or false: Any two oriented linear spaces of the same dimension are isomorphic.

Answer: True when their common dimension is positive; but the spaces $\Diamond$ and $-\Diamond$ are not isomorphic. There is precisely one bijection between them, the zero map; but that map does not preserve orientation. This is one aspect of what it means for the orientation of a zero-dimensional oriented space to be an intrinsic aspect of that space.

Exercise 5-2 (For people who know about alternating tensors) Let M be a real linear space of finite dimension $m := \dim(M)$. The m^{th} *exterior power* of M is the space $\bigwedge^m M$ of all alternating, m-contravariant tensors on M, and we have

$\dim(\bigwedge^m M) = \binom{m}{m} = 1$. One simple example of such a tensor is $\mu_1 \wedge \ldots \wedge \mu_m$, where $(\mu_1, \ldots, \mu_m)$ is a basis for M. Show that orienting M is equivalent to distinguishing one of the two rays leaving the origin in $\bigwedge^m M$ as the positive ray. In the special case where M is zero-dimensional, note that $\bigwedge^m M = \mathbb{R}$; and thus, in this framework also, the two possible orientations of a zero-dimensional space are intrinsically distinguishable.

Comment: People integrating over manifolds deal with a related space, the space $\bigwedge^m(M^*)$ of all alternating, m-covariant tensors on M. If $(d\mu_1, \ldots, d\mu_m)$ is the basis for the dual space M^* that is dual to the basis $(\mu_1, \ldots, \mu_m)$ for M, then $d\mu_1 \wedge \ldots \wedge d\mu_m$ is a simple example of an alternating, m-covariant tensor, sometimes called a *volume form*. Orienting a linear space M is equivalent to orienting its dual M^*; but we won't have any occasion to exploit dual spaces in this monograph.

If $f: K \to L$ is a linear bijection between two oriented linear spaces K and L, then the orientations on K and L are related by f. We denote by $f(K)$ the space L, but with the orientation carried forward from K via f. That is, we have $f(K) = L$ or $f(K) = -L$, according as the bijection $f: K \to L$ preserves or reverses orientation. So we always have $f(K) = \operatorname{sgn}(f)L$.

More generally, whenever $f: K \to L$ is injective, we use $f(K)$ to denote $\operatorname{Im}(f)$ with the orientation carried forward from K. Note that every injective map f preserves orientation when f is interpreted as a bijection from K to $f(K)$.

Still more generally, if $f: K \to L$ is any linear map, the expression $f(U)$ makes sense as an oriented linear subspace of L whenever U is an oriented subspace of K with $U \cap \operatorname{Ker}(f) = \{0\}$. We use $f(U)$ to denote the image of the subspace U under f, with the orientation carried forward from U.

Exercise 5-3 If M is an oriented linear space of dimension m, let $\neg: M \to M$ denote the negating map, the map that takes each vector μ in M to its negative: $\neg(\mu) := -\mu$. Show that $\neg(M) = (-1)^m M$; that is, negating all m coordinates of a space reverses the orientation of that space precisely when m is odd. It follows that we must distinguish carefully between $\neg(M)$ and $-M$.

5.2 Direct sums and direct products

Suppose that K and L are linearly independent subspaces of some larger space. Given a signed, ordered basis $((\kappa_1, \ldots, \kappa_k), \beta)$ for K and one $((\lambda_1, \ldots, \lambda_l), \gamma)$ for L, the obvious way to orient the direct sum $K \oplus L$ is with the signed, ordered basis

$$((\kappa_1, \ldots, \kappa_k, \lambda_1, \ldots, \lambda_l),\ \beta\gamma).$$

That is, we concatenate the two bases with the basis of K going first, and we multiply the two signs. This rule combines the given orientations on K and L

to produce an orientation on the direct sum $K \oplus L$, an orientation that does not depend upon which signed, ordered bases we choose for K and L. We shall refer to this rule for orienting a direct sum of oriented subspaces as the *Concatenate Rule*. Roughly speaking, we form a positive basis for $K \oplus L$ by concatenating positive bases for K and for L, in that order — the only roughness about that recipe being that a negatively oriented, zero-dimensional space doesn't have any positive bases.

The Concatenate Rule behaves well with respect to reversing orientation: If K and L are linearly independent, oriented subspaces of some larger space, we have

$$(-K) \oplus L = K \oplus (-L) = -(K \oplus L).$$

But the order of the direct summands is important; we have

$$K \oplus L = (-1)^{kl}(L \oplus K).$$

Thus, which basis goes first, K or L, in building an oriented basis for the direct sum $K \oplus L$, makes a difference precisely when both K and L are odd-dimensional.

Suppose now that K and L are any two oriented spaces, not subspaces of some common, larger space. We shall use the Concatenate Rule also to orient the direct product $K \times L$ in the obvious way. That is, we decompose the direct product $K \times L$ as the direct sum $K \times L = (K, 0) \oplus (0, L)$ of the K-axis, the set $(K, 0) := \{(\kappa, 0) \mid \kappa \in K\}$, and the analogous L-axis, and we then apply the Concatenate Rule for direct sums.

Warning: This is the first situation that we have come across, in computing either direct products or fiber products, where the order of the factors is critical. Given two linear spaces K_1 and K_2, we have so far been thinking of the direct products $K_1 \times K_2$ and $K_2 \times K_1$ as two ways of viewing the same space: the space of all functions κ defined on the index set $\{1, 2\}$ for which κ_1 lies in K_1 and κ_2 lies in K_2. That is, we have been exploiting the ordering on the index set $\{1, 2\}$ only in unimportant ways, such as deciding which factor to write to the left of the "$\times$". But the Concatenate Rule exploits that ordering in a critical way, to determine which basis goes first. If we use the swapping map $(\kappa, \lambda) \mapsto (\lambda, \kappa)$ to identify $K \times L$ with $L \times K$, we have, as oriented spaces:

$$K \times L = (-1)^{kl}(L \times K).$$

Thus, when taking direct products of odd-dimensional oriented linear spaces, the order of the factors is critical.

Exercise 5-4 Let ξ and η be unit vectors in the directions of the positive x and y axes, so that the linear space $\langle \xi, \eta \rangle$ is the (x, y)-plane with its standard orientation. Mathematicians generally draw the plane $\langle \xi, \eta \rangle$ so that the 90-degree rotation that takes ξ to η is counterclockwise. Adopting both that convention and the Concatenate Rule, consider Figures 1.7 and 1.11. Which of them illustrates the state space $A \times B$ and which illustrates $B \times A = -(A \times B)$?

Answer: Figure 1.7 illustrates $A \times B$.

5.3 Quotient spaces

Let L be a linear subspace of a linear space M, and consider the quotient space M/L. Given any basis for L, we can extend that basis into a basis for M. And the vectors in that extension — more properly, their equivalence classes modulo L — constitute a basis for the quotient space M/L. If we start with a positive basis for L and we extend it so as to form a positive basis for M, we can orient the quotient space M/L by taking the extending vectors, in that same order, as a positive basis for M/L. This is roughly the same as using the Concatenate Rule on the direct product $L \times (M/L)$ (which is isomorphic to M, but not canonically). Unfortunately, it would be equally reasonable to concatenate in the other order, as in the direct product $(M/L) \times L$, putting the extending vectors first. There is no compelling reason to prefer one order over the other.

The same need for a choice arises if we deal with complements, which are sometimes more convenient than quotient spaces. Note that orienting the quotient space M/L is equivalent to orienting any (and hence every) complement C of L in M. Indeed, if we extend a basis for L into a basis for M, the linear span of the extending vectors will be a complement C of L in M. Thus, instead of relating the orientations on M, L, and M/L, we can relate the orientations on M, L, and C, using one of the direct sums $M = C \oplus L$ or $M = L \oplus C$. But we still have to choose one of those two.

Returning from complements to quotient spaces, note that the unfortunate need to choose between $(M/L) \times L$ and $L \times (M/L)$ arises only when L and M/L might both be odd-dimensional. Suppose, for example, that $L = \pm M$. The quotient space M/L is then zero-dimensional. Since the integer 0 is definitely even, the order of the two factors does not matter in this case; the orientations on L and on M determine the orientation on the zero-dimensional space M/L uniquely. In particular, we have $M/L = \lozenge$ when $L = M$, and we have $M/L = -\lozenge$ when $L = -M$.

It is sometimes convenient to use the two linear spaces $\lozenge$ and $-\lozenge$, rather than the two integers 1 and -1, to encode the two possible signs of a bijection. If the map $f\colon K \to L$ is a linear bijection, we can then denote $\operatorname{sgn}(f)$ simply as the quotient space $f(K)/L$.

Exercise 5-5 Let $f\colon K \to L$ be any linear map and let C_1 and C_2 be two oriented complements of $\operatorname{Ker}(f)$ in K. Fixing arbitrary orientations on K, $\operatorname{Ker}(f)$, and $\operatorname{Im}(f)$, we have $\operatorname{Ker}(f) \oplus C_1 = \pm K$ and $f(C_1) = \pm \operatorname{Im}(f)$, and similarly for C_2, though possibly with different signs. Show that $(\operatorname{Ker}(f) \oplus C_1)/(\operatorname{Ker}(f) \oplus C_2) = f(C_1)/f(C_2)$; that is, either both quotients are $+\lozenge$ or both are $-\lozenge$.

Hint: Let $q\colon K \to K/\operatorname{Ker}(f)$ be the quotient map. We have $q(C_1) = \pm(K/\operatorname{Ker}(f))$, and similarly for C_2. Show that both of the quotients above are equal to $q(C_1)/q(C_2)$.

5.4 Matrices of maps

We are going to be dealing with direct products of linear spaces and with linear maps that have such direct products as their domains or codomains. In doing computations with such a map, it can be helpful to represent it as a matrix whose entries are themselves maps, rather than scalars. Indeed, the theory of matrices can be generalized even further by letting the matrix entries be morphisms from any additive category [12]. But linear maps are general enough for us. Let's settle on a consistent set of conventions for that case.

Suppose that A, B, K, and L are linear spaces. Let $f_{K\leftarrow A}$ be some linear map $f_{K\leftarrow A}\colon A \to K$, and similarly for $f_{L\leftarrow A}$, $f_{K\leftarrow B}$, and $f_{L\leftarrow B}$. (We write these subscripts in the order "codomain $\leftarrow$ domain" rather than "domain $\to$ codomain" because, in this monograph, we are writing our maps and their matrices as prefix operators.) We can combine these maps in various ways.

- The operation of direct product is the most general way to combine two maps, and it does not require any special properties of the maps being combined. For example, the direct product $f_{K\leftarrow A} \times f_{L\leftarrow B}$ is the linear map from $A \times B$ to $K \times L$ given by

  $$(f_{K\leftarrow A} \times f_{L\leftarrow B})(\alpha, \beta) := (f_{K\leftarrow A}(\alpha),\ f_{L\leftarrow B}(\beta)).$$

  Note that this same combining operation is often described as a direct sum. If A and B are linearly independent subspaces of some larger domain space, while K and L are linearly independent subspaces of some larger codomain space, then the map $f_{K\leftarrow A} \times f_{L\leftarrow B}$ can also be described as the direct sum $f_{K\leftarrow A} \oplus f_{L\leftarrow B}\colon A \oplus B \to K \oplus L$.

- Given two linear maps defined on the same domain, we can combine them in a tighter way by sharing the argument. For example, the two maps $f_{K\leftarrow A}$ and $f_{L\leftarrow A}$ share the domain A; we can combine them to get the map $(f_{K\leftarrow A},\ f_{L\leftarrow A})\colon A \to K \times L$ given by

  $$(f_{K\leftarrow A},\ f_{L\leftarrow A})(\alpha) := (f_{K\leftarrow A}(\alpha),\ f_{L\leftarrow A}(\alpha)).$$

- Given two linear maps that share the same codomain, we can combine them in a tighter way by adding the results. For example, we combine the maps $f_{K\leftarrow A}$ and $f_{K\leftarrow B}$, which share the codomain K, to get the map $f_{K\leftarrow A} \mathbin{+\!\!\!+} f_{K\leftarrow B}\colon A \times B \to K$ given by

  $$(f_{K\leftarrow A} \mathbin{+\!\!\!+} f_{K\leftarrow B})(\alpha, \beta) := f_{K\leftarrow A}(\alpha) + f_{K\leftarrow B}(\beta).$$

  The exotic symbol "$+\!\!\!+$" warns that this binary operator is something fancier than simple addition.

Suppose that we start off with four maps $f_{K\leftarrow A}$, $f_{K\leftarrow B}$, $f_{L\leftarrow A}$, and $f_{L\leftarrow B}$. We can assemble those four into a single map $\Phi\colon A\times B \to K\times L$ by putting the latter two combining operations together in several different ways: We can either set $\Phi := (f_{K\leftarrow A},\ f_{L\leftarrow A}) \mathbin{\#} (f_{K\leftarrow B},\ f_{L\leftarrow B})$ or we can set $\Phi := (f_{K\leftarrow A} \mathbin{\#} f_{K\leftarrow B},\ f_{L\leftarrow A} \mathbin{\#} f_{L\leftarrow B})$. Either way, we end up with the same map

$$\Phi(\alpha,\beta) = \bigl(f_{K\leftarrow A}(\alpha) + f_{K\leftarrow B}(\beta),\ f_{L\leftarrow A}(\alpha) + f_{L\leftarrow B}(\beta)\bigr).$$

Perhaps the clearest way to write this map Φ is as a 2-by-2 matrix. Adopting the conventions for matrices in which matrices are prefix operators, we write the application $\Phi(\alpha,\beta) = (\kappa,\lambda)$ as

$$\begin{pmatrix} f_{K\leftarrow A} & f_{K\leftarrow B} \\ f_{L\leftarrow A} & f_{L\leftarrow B} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \kappa \\ \lambda \end{pmatrix}.$$

Note that, under these conventions, the maps with the same domain form a column of the matrix, while the maps with the same codomain form a row. In particular, using $[g]$ to denote the matrix of the linear map g, we have

$$[(f_{K\leftarrow A},\ f_{L\leftarrow A})] = \begin{pmatrix} [f_{K\leftarrow A}] \\ [f_{L\leftarrow A}] \end{pmatrix}, \tag{5-6}$$

while

$$[f_{K\leftarrow A} \mathbin{\#} f_{K\leftarrow B}] = \begin{pmatrix} [f_{K\leftarrow A}] & [f_{K\leftarrow B}] \end{pmatrix}. \tag{5-7}$$

Equation 5-6 combines two maps with the same domain by pasting their matrices together, one above the other. Equation 5-7 combines two maps with the same codomain by pasting their matrices together, side by side.

Exercise 5-8 How does the direct-product operator combine matrices? That is, what is the analog of Equations 5-6 and 5-7 for the map $f_{K\leftarrow A} \times f_{L\leftarrow B}$?

Answer: The direct-product operator assembles blocks down the diagonal:

$$[f_{K\leftarrow A} \times f_{L\leftarrow B}] = \begin{pmatrix} [f_{K\leftarrow A}] & 0 \\ 0 & [f_{K\leftarrow B}] \end{pmatrix}$$

5.5 Elementary operations on matrices of maps

Suppose that we are dealing with a linear map whose domain and codomain are direct products of smaller spaces. We just adopted conventions for representing such a map as a matrix whose entries are themselves maps. One advantage of this representation is that we can perform elementary row and column operations on the resulting matrix; for example, we can add a multiple of any row to any other row without affecting the overall determinant. This is just what happens in standard linear algebra, where the entries in the matrices are scalars. When the matrix elements are maps, however, things are more subtle, since the composition operator on functions is not commutative. Thus, when we "multiply" the maps in one row by a constant map c, we have to compose by c on the proper side — as it turns out, for a row operation, we compose with c on the left. We here review the theory of elementary row and column operations to see how it extends to the noncommutative case.

To see the general pattern, it suffices to consider a linear map $\Phi\colon A \times B \to K \times L$ whose matrix representation

$$[\Phi] = \begin{pmatrix} f_{K\leftarrow A} & f_{K\leftarrow B} \\ f_{L\leftarrow A} & f_{L\leftarrow B} \end{pmatrix}$$

is 2-by-2. We first consider performing an elementary column operation on the matrix $[\Phi]$, say adding some multiple of the second column to the first column. To achieve that goal, we introduce the map $\Psi\colon A \times B \to A \times B$ whose matrix is

$$[\Psi] = \begin{pmatrix} 1_{A\leftarrow A} & 0_{A\leftarrow B} \\ g_{B\leftarrow A} & 1_{B\leftarrow B} \end{pmatrix}.$$

The lower-left entry here is an arbitrary linear function $g_{B\leftarrow A}\colon A \to B$, which is going to play the role of the arbitrary scalar by which we multiply in a standard column operation. Performing the column operation involves replacing the matrix $[\Phi]$ with the matrix product $[\Phi][\Psi] = [\Phi \circ \Psi]$, as follows:

$$[\Phi][\Psi] = \begin{pmatrix} f_{K\leftarrow A} & f_{K\leftarrow B} \\ f_{L\leftarrow A} & f_{L\leftarrow B} \end{pmatrix} \begin{pmatrix} 1_{A\leftarrow A} & 0_{A\leftarrow B} \\ g_{B\leftarrow A} & 1_{B\leftarrow B} \end{pmatrix} = \begin{pmatrix} f_{K\leftarrow A} + f_{K\leftarrow B}\, g_{B\leftarrow A} & f_{K\leftarrow B} \\ f_{L\leftarrow A} + f_{L\leftarrow B}\, g_{B\leftarrow A} & f_{L\leftarrow B} \end{pmatrix}.$$

To move from the matrix $[\Phi]$ to $[\Phi][\Psi]$, we compose the second column on the right by the function $g_{B\leftarrow A}$ and add the result to the first column. That process constitutes an *elementary column operation*. Note that the matrix $[\Psi]$ is lower-triangular with 1's on the diagonal, no matter what bases we choose for the spaces A and B. Hence, we have $\det([\Psi]) = 1$, from which it follows that $\det([\Phi][\Psi]) = \det([\Phi])$. Thus, performing an elementary column operation does not affect the determinant.

Dually, we can also perform *elementary row operations* on a matrix without affecting its determinant: We compose some row on the left by an appropriate function and add the result to some other row. For example, for any linear map $h_{L\leftarrow K}\colon K \to L$, we can transform $[\Phi]$ like this:

$$\begin{pmatrix} 1_{K\leftarrow K} & 0_{K\leftarrow L} \\ h_{L\leftarrow K} & 1_{L\leftarrow L} \end{pmatrix} \begin{pmatrix} f_{K\leftarrow A} & f_{K\leftarrow B} \\ f_{L\leftarrow A} & f_{L\leftarrow B} \end{pmatrix} = \begin{pmatrix} f_{K\leftarrow A} & f_{K\leftarrow B} \\ h_{L\leftarrow K}\, f_{K\leftarrow A} + f_{L\leftarrow A} & h_{L\leftarrow K}\, f_{K\leftarrow B} + f_{L\leftarrow B} \end{pmatrix}.$$

These elementary row and column operations will be one of our tools when we need to prove things about bijections between spaces that are direct products.

5.6 Orienting a smooth manifold

Roughly speaking, we orient a smooth manifold by orienting each of its tangent spaces, with the proviso that the orientations on the tangent spaces must be chosen in a locally consistent manner. But there are some subtleties. First, there are some manifolds that simply cannot be oriented; their global geometry does not permit the required local consistency to hold everywhere. Second, for those manifolds that can be oriented, the number of possible orientations depends upon the number of connected components: An orientable manifold with k connected components has precisely 2^k possible orientations.

In more detail, let $\mathbf{M}$ be a smooth manifold, of dimension m. At each point $\mathbf{m}$ in $\mathbf{M}$, the tangent space $T_{\mathbf{m}}\mathbf{M}$ is a linear m-space. Suppose that we choose one of its two possible orientations, for each point $\mathbf{m}$ in $\mathbf{M}$, in some arbitrary way. Let's refer to the result as a *pseudo-orientation* of the manifold $\mathbf{M}$.

In order for a pseudo-orientation of a manifold $\mathbf{M}$ to qualify as an orientation, it must be locally consistent; what does that mean? Consider a chart $\varphi\colon U \to \mathbb{R}^m$ on some open subset U of $\mathbf{M}$. At each point $\mathbf{m}$ in the domain U of the chart φ, we can use the tangents to the m coordinate axes of $\mathbb{R}^m$, in numeric order, to get a reference orientation on the tangent space $T_{\mathbf{m}}\mathbf{M}$. We say that a pseudo-orientation of $\mathbf{M}$ is *consistent* with the chart $\varphi\colon U \to \mathbb{R}^m$ when, for all points $\mathbf{m}$ in U, the pseudo-orientation orients the tangent space $T_{\mathbf{m}}\mathbf{M}$ in agreement with the reference orientation provided by φ. A pseudo-orientation of $\mathbf{M}$ is *locally consistent at a point* $\mathbf{m}$ in $\mathbf{M}$ when there is some chart on some neighborhood of $\mathbf{m}$ with which the pseudo-orientation is consistent. Finally, we say that a pseudo-orientation of $\mathbf{M}$ is *locally consistent* when it is locally consistent at every point $\mathbf{m}$ in $\mathbf{M}$.

There are manifolds, called *non-orientable*, on which this local consistency simply cannot be achieved. The Möbius strip and the Klein bottle are simple examples of non-orientable 2-manifolds. Fortunately, the direct products and the fiber products of orientable manifolds are always orientable; we shall prove that result for direct products in a moment and for fiber products in Section 7.3. So, if we start with manifolds that are orientable and we build new manifolds only by taking direct and fiber products, the problem of non-orientability never arises.

Each connected component of an orientable manifold is itself orientable and has precisely 2 possible orientations. A manifold with k connected components, thus, has a total of 2^k possible orientations. The case $k = 0$ is worth noting. Since $2^0 = 1$, an empty manifold has only one possible orientation. So orienting an empty manifold is trivial; there are no choices involved, because there are no tangent spaces to orient.

5.7 Orienting direct products of manifolds

We saw, in Section 5.2, that the Concatenate Rule gives us a way to orient the direct products of oriented linear spaces. By using the Concatenate Rule as a linear-space subroutine, independently at each point, we can get an analogous rule for orienting the direct products of oriented manifolds.

Consider some direct product $\mathbf{D} = \mathbf{A}_1 \times \cdots \times \mathbf{A}_n$ of oriented manifolds, and suppose that the factor manifolds $\mathbf{A}_1$ through $\mathbf{A}_n$ have been given to us in a specified order. Standard theory tells us that, at each point, the tangent space to the direct product is canonically isomorphic to the direct product of the tangent spaces; that is, for any point $\mathbf{d} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ in $\mathbf{D}$, we have

$$T_{\mathbf{d}}\mathbf{D} \cong (T_{\mathbf{a}_1}\mathbf{A}_1) \times \cdots \times (T_{\mathbf{a}_n}\mathbf{A}_n).$$

So, at each point $\mathbf{d}$ in $\mathbf{D}$, we can produce a basis for the tangent space $T_{\mathbf{d}}\mathbf{D}$ by concatenating, in the specified order, positive bases for the factor tangent spaces $T_{\mathbf{a}_1}\mathbf{A}_1$ through $T_{\mathbf{a}_n}\mathbf{A}_n$. The *Concatenate Rule for Manifolds* orients the tangent space $T_{\mathbf{d}}\mathbf{D}$ so that a basis constructed in this way is positive.

Why does this rule give an orientation of the manifold $\mathbf{D}$, and not just a pseudo-orientation? That is, why is the Concatenate Rule locally consistent, say at the point $\mathbf{d} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$? For each i from 1 to n, by the local consistency in the factor manifold $\mathbf{A}_i$, we can find a chart $\varphi_i \colon U_i \to \mathbb{R}^{a_i}$ on some neighborhood U_i of $\mathbf{a}_i$ with which the given orientation on $\mathbf{A}_i$ is consistent. The direct product $U_1 \times \cdots \times U_n$ is then an open neighborhood of the point $\mathbf{d}$ in $\mathbf{D}$, and we can define a chart $\psi \colon U_1 \times \cdots \times U_n \to \mathbb{R}^d$ on that neighborhood by concatenating sequences of coordinates, setting

$$\psi(\mathbf{a}'_1, \ldots, \mathbf{a}'_n) := \bigl(\varphi_1(\mathbf{a}'_1), \ldots, \varphi_n(\mathbf{a}'_n)\bigr).$$

The pseudo-orientation specified by the Concatenate Rule is consistent with this chart ψ over the entire neighborhood $U_1 \times \cdots \times U_n$ of $\mathbf{d}$ and is hence locally consistent. Thus, the Concatenate Rule does indeed orient the direct products of oriented manifolds. It follows that such direct products are always orientable, as we claimed in Section 5.6.

Chapter 6

Stability

In this monograph, we are going to develop a compelling rule for orienting the transverse fiber products of oriented linear spaces. Given such a rule for linear spaces, we then intend to produce an analogous rule for manifolds by applying our linear-space rule, as a subroutine, to each tangent space independently. That approach is an obvious one, and the side condition of transversality is no obstacle. But we must require a certain continuity property of our linear-space rule, lest we end up with pseudo-orientations of some fiber product manifolds that are not locally consistent. Thinking back to Alice and Bob, the Greedy-Alice Rule in Figure 1.8 gives an example of what could go wrong were our linear-space rule to lack the required continuity. In this chapter, we define the required continuity property — which we christen “stability”.

How does this subroutine stuff work, in detail? Let $\mathbf{S}$ and $\mathbf{A}_1$ through $\mathbf{A}_n$ be smooth, oriented manifolds. Suppose that $\mathbf{f}_i\colon \mathbf{A}_i \to \mathbf{S}$ is a smooth map, for i from 1 to n, and that the maps $(\mathbf{f}_1, \ldots, \mathbf{f}_n)$ are transverse. By Proposition 4-12, the fiber product $\mathbf{P} := \mathbf{A}_1 \times_{\mathbf{S}} \cdots \times_{\mathbf{S}} \mathbf{A}_n$ is again a smooth manifold. Let's view that fiber product $\mathbf{P}$ as a subset — actually, a submanifold — of the direct product $\mathbf{A}_1 \times \cdots \times \mathbf{A}_n$, and let $\mathbf{p} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ be a point in $\mathbf{P}$. So there exists a point $\mathbf{s}$ in $\mathbf{S}$ with $\mathbf{f}_i(\mathbf{a}_i) = \mathbf{s}$, for i from 1 to n. By standard theory once again, the tangent space to the fiber product is naturally isomorphic to the fiber product of the tangent spaces; that is, we have

$$T_{\mathbf{p}}\mathbf{P} \cong (T_{\mathbf{a}_1}\mathbf{A}_1)[T_{\mathbf{a}_1}\mathbf{f}_1] \times_{T_{\mathbf{s}}\mathbf{S}} \cdots \times_{T_{\mathbf{s}}\mathbf{S}} [T_{\mathbf{a}_n}\mathbf{f}_n](T_{\mathbf{a}_n}\mathbf{A}_n).$$

We shall often abbreviate this as $P = A_1[f_1] \times_S \cdots \times_S [f_n]A_n$. We can orient the tangent space P on the left by using our assumed linear-space rule, as a subroutine, to orient the fiber product $A_1 \times_S \cdots \times_S A_n$ on the right. The linear spaces A_1 through A_n and S come to us oriented, because they are tangent spaces to the oriented manifolds $\mathbf{A}_1$ through $\mathbf{A}_n$ and $\mathbf{S}$. Furthermore, from Definition 4-11, what it means for the smooth maps $(\mathbf{f}_i)$ to be transverse is that their differentials (f_i) are transverse linear maps at every point $\mathbf{p} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ in the fiber product $\mathbf{P}$. So we are in fine shape to invoke our subroutine, that is, to orient the tangent space $P = T_{\mathbf{p}}\mathbf{P}$ as a transverse fiber product of oriented linear spaces.

But what about local consistency? As the point $\mathbf{p} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ moves around in the fiber product $\mathbf{P}$, the orientations on the varying tangent spaces S and A_1 through A_n will be locally consistent. Also, the factor maps (f_i) will vary continuously and will remain transverse. Does this ensure that the orientations that we assign to the tangent spaces $P = A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ by repeatedly calling our linear-space subroutine will also be locally consistent?

Fiber products of smooth manifolds are subtle to think about, since so many things vary simultaneously. As the point $\mathbf{p} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ varies over the fiber product $\mathbf{P} = \mathbf{A}_1 \times_{\mathbf{S}} \cdots \times_{\mathbf{S}} \mathbf{A}_n$, the tangent spaces $A_i := T_{\mathbf{a}_i}\mathbf{A}_i$ and $S := T_{\mathbf{s}}\mathbf{S}$ vary, along with the factor maps $f_i \colon A_i \to S$. It would be simpler if the linear spaces involved stayed fixed and only the maps relating those spaces varied — for example, if there were a single, fixed linear space A_1, rather than separate linear spaces associated with each of the points $\mathbf{a}_1$ in $\mathbf{A}_1$. We show in Section 6.3 that we can convert to that simpler situation by restricting the point $\mathbf{p}$ to move only in a small region, a region small enough so that the point $\mathbf{a}_i$ remains within a local coordinate system on the manifold $\mathbf{A}_i$, for i from 1 to n, while the point $\mathbf{s}$ remains within a local coordinate system on the manifold $\mathbf{S}$. In Sections 6.1 and 6.2, we study the simpler situation in which the spaces stay fixed.

Remark: There are well-developed mathematical techniques for dealing with all of the tangent spaces to a smooth manifold at once, without limiting ourselves to a portion of the manifold that has a local coordinate system. All of the tangent spaces together form the *tangent bundle* of the manifold. Tangent bundles are an important example of *vector bundles*, which, in turn, are an important example of *fiber bundles* — the area of mathematics where the name "fiber product" came from. But we can do what we need to do in this monograph working locally, without recourse to tangent bundles.

6.1 Varying only the factor maps

In this section and the next, we take the linear spaces A_1 through A_n and S to be fixed, and we allow only the factor maps $f_i \colon A_i \to S$ to vary. What does the fiber-product operator itself correspond to, in this simpler situation?

Definition 6-1 If A and B are linear spaces, let $\mathrm{Lin}(A, B)$ denote the set of all linear maps from A to B. Each such linear map corresponds to a matrix with a columns and b rows, and we equip the space $\mathrm{Lin}(A, B)$ with the topology of $\mathbb{R}^{ab}$.

Definition 6-2 If A_1 through A_n and S are linear spaces, let $\mathrm{Trans}(A_1, \ldots, A_n; S)$ denote that subset of the direct product $\mathrm{Lin}(A_1, S) \times \cdots \times \mathrm{Lin}(A_n, S)$ consisting of n-tuples of maps $(f_1, \ldots, f_n)$ that are transverse.

Given any n-tuple of maps $(f_1, \ldots, f_n)$ in $\mathrm{Trans}(A_1, \ldots, A_n; S)$, we can form the fiber product

$$P(f_1, \ldots, f_n) := A_1[f_1] \times_S \cdots \times_S [f_n]A_n.$$

This fiber product $P(f_1, \ldots, f_n)$ is some linear subspace of the direct product $D := A_1 \times \cdots \times A_n$, and its dimension is $p := a_1 + \cdots + a_n - (n-1)s$. The set of all linear subspaces of a fixed space D that have a fixed dimension p is a well-known manifold called a *Grassmannian*.

Definition 6-3 If K is a linear space of dimension k and we have $0 \leq d \leq k$, the set of all linear subspaces of K of dimension d has a natural structure as a smooth manifold $\mathcal{G}(d, K)$, called the *Grassmannian of d-dimensional subspaces of K*. We could express any particular basis for such a subspace as a k-by-d matrix; but multiplying this matrix on the right by any invertible d-by-d matrix would produce another basis for the same d-dimensional subspace. Thus, the dimension of the manifold $\mathcal{G}(d, K)$ is $kd - d^2 = d(k-d)$.

For example, when $d = 1$, consider all of the lines through the origin of the linear space K. Those lines form a projective space of dimension $d(k-d) = k-1$; so the Grassmann manifold $\mathcal{G}(1, K)$ is simply projective $(k-1)$-space.

The fiber-product operator P discussed above can hence be viewed as a map

$$P\colon \operatorname{Trans}(A_1, \ldots, A_n; S) \to \mathcal{G}(p, D).$$

Note that this map P is continuous: Small changes to the factor maps f_1 through f_n cause only small changes in the resulting linear subspace $P(f_1, \ldots, f_n)$.

Exercise 6-4 Let A, B, and S be linear spaces of dimensions a, b, and s with $a + b \geq s$, and consider the fiber-product operator

$$P_{(A,B;S)}\colon \operatorname{Trans}(A, B; S) \to \mathcal{G}(a + b - s, A \times B).$$

What is the dimension of the domain of $P_{(A,B;S)}$? Of its codomain? Account for the discrepancy between the two dimensions.

Answer: The product space $\operatorname{Lin}(A, S) \times \operatorname{Lin}(B, S)$ has dimension $as + bs = (a+b)s$. When $a + b \geq s$, most pairs of maps (f, g) in $\operatorname{Lin}(A, S) \times \operatorname{Lin}(B, S)$ are transverse, so the subset $\operatorname{Trans}(A, B; S)$ has that same dimension. But the Grassmann manifold $\mathcal{G}(a + b - s, A \times B)$ has dimension only

$$(a + b - s)\big((a + b) - (a + b - s)\big) = (a + b - s)s.$$

The former exceeds the latter by s^2. This discrepancy arises because, for any invertible linear map $e\colon S \to S$, the pair of factor maps $(e \circ f, e \circ g)$ gives the same fiber product as the pair (f, g); and there are s^2 degrees of freedom in the choice of the map e.

Exercise 6-5 The previous exercise works out quite neatly, in part because the binary fiber-product map $P_{(A,B;S)}$ is surjective; that is, every linear subspace of $A \times B$ of dimension $a + b - s$ can arise as a fiber product $A[f] \times_S [g]B$. Once n exceeds 2, the n-ary case is not so simple. For example, consider ternary fiber products in which A, B, and C are copies of the plane $\mathbb{R}^2$, while S is the line $\mathbb{R}$. So a transverse fiber product of the form $A[f] \times_S [g]B \times_S [h]C$ is 4-dimensional. What is the dimension of the domain $\mathrm{Trans}(A, B, C; S)$ of the fiber-product map $P_{(A,B,C;S)}$? Of its codomain $\mathcal{G}(4, A \times B \times C)$? Account for the discrepancy.

Answer: The factor map $f : \mathbb{R}^2 \to \mathbb{R}$ has two degrees of freedom in it, and the same for g and h; so the triple of maps (f, g, h) has six degrees of freedom. For generic choices of those three maps, the difference map $\Delta\colon A \times B \times C \to S^2$ is surjective; so the transverse subset $\mathrm{Trans}(A, B, C; S)$ is also 6-dimensional. But one of those dimensions is crushed out by the fiber-product map $P_{(A,B,C;S)}$, since the triple of maps $(e \circ f, e \circ g, e \circ h)$ gives the same fiber product as (f, g, h), for any invertible linear map $e\colon \mathbb{R} \to \mathbb{R}$. So we expect the image of the fiber-product map $P_{(A,B,C;S)}$ to be 5-dimensional.

As for the full Grassmann manifold $\mathcal{G}(4, A \times B \times C)$ of all 4-dimensional linear subspaces of a linear 6-space, it has dimension $4(6 - 4) = 8$. So there remains a discrepancy of $8 - 5 = 3$.

To see why a typical 4-dimensional subspace of $A \times B \times C$ can't arise as a fiber product, consider the factor map $f : A \to S$. Since $\dim(A) = 2 > 1 = \dim(S)$, there must exist a nonzero vector α in A with $f(\alpha) = 0$. And the vector $(\alpha, 0, 0)$ in $A \times B \times C$ then belongs to the fiber product. Thus, every 4-dimensional subspace of the 6-space $A \times B \times C$ that arises as a fiber product has a nontrivial intersection with the fixed subspace $A \times 0 \times 0$. But a typical 4-subspace of 6-space intersects a fixed 2-subspace only in the origin. It costs one degree of freedom to insist that our 4-subspace intersect $A \times 0 \times 0$ in an entire line. It costs two additional degrees of freedom to insist that it similarly intersect $0 \times B \times 0$ and $0 \times 0 \times C$.

6.2 Stability defined

We continue to restrict ourselves to the simpler situation in which the linear spaces A_1 through A_n and S stay fixed, while only the factor maps $f_i\colon A_i \to S$ are allowed to vary. Our next goal is to define what it means, in that simpler situation, for an orientation rule to be *stable*.

Each d-dimensional linear subspace of a k-dimensional linear space K can be given two possible orientations. Suppose that we distinguish between them. If we assemble together all of the resulting oriented d-subspaces of K, the result is a smooth manifold that is, in some sense, "twice as big" as the standard Grassmannian $\mathcal{G}(d, K)$. This larger manifold is called the *Grassmannian of oriented d-dimensional subspaces of K*, and we shall write it $\vec{\mathcal{G}}(d, K)$. For example, in the case $d = 1$, the Grassmannian $\vec{\mathcal{G}}(1, K)$ consists of all oriented lines through the origin of K, which form simply the sphere $\mathbb{S}^{k-1}$.

Forgetting about the orientation of an oriented d-subspace of K leaves us with an unoriented d-subspace. So we have a natural map from the big Grassmannian to the standard one, from $\vec{\mathcal{G}}(d, K)$ to $\mathcal{G}(d, K)$. This map is precisely two-to-one. In fact, every neighborhood in $\mathcal{G}(d, K)$ that is small enough not to include both a d-subspace and its negative has two isomorphic pre-images in $\vec{\mathcal{G}}(d, K)$; one says that $\vec{\mathcal{G}}(d, K)$ is a *double cover* of $\mathcal{G}(d, K)$. For example, when $d = 1$, the $(k-1)$-sphere $\mathbb{S}^{k-1}$ is a double cover of projective $(k-1)$-space.

Let S and A_1 through A_n be fixed linear spaces. We saw in Section 6.1 that taking fiber products gives us a continuous map

$$P\colon \operatorname{Trans}(A_1, \ldots, A_n; S) \to \mathcal{G}(p, A_1 \times \cdots \times A_n),$$

where $p := a_1 + \cdots + a_n - (n-1)s$ is the dimension of such a transverse fiber product. Suppose now that the spaces S and A_1 through A_n are oriented and that we have some rule in mind for orienting the resulting fiber products. By exploiting that rule, we get a map

$$\vec{P}\colon \operatorname{Trans}(A_1, \ldots, A_n; S) \to \vec{\mathcal{G}}(p, A_1 \times \cdots \times A_n).$$

We are particularly interested in those orientation rules for which this more refined map $\vec{P}$ is also continuous, so that small changes to the factor maps f_1 through f_n cause the oriented linear subspace $\vec{P}(f_1, \ldots, f_n)$ to slew around slightly, but without any sudden reversals of its orientation.

Definition 6-6 A rule for orienting the transverse fiber products of oriented linear spaces is *stable* when, for all oriented linear spaces A_1 through A_n and S, the map that it defines from

$$\operatorname{Trans}(A_1, \ldots, A_n; S) \to \vec{\mathcal{G}}(a_1 + \cdots + a_n - (n-1)s, A_1 \times \cdots \times A_n)$$

taking the transverse n-tuple of maps $(f_1, \ldots, f_n)$ to the oriented fiber product $A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ is continuous.

Definition 6-6 is global and elegant; but we couldn't actually use it in a proof unless we took the time to define the Grassmann manifold precisely — that is, to construct explicit charts for it. Instead, we shall view Definition 6-6 as motivation and define stability in the following more concrete and local way. Readers who know about Grassmann manifolds will be able to verify that the two definitions are equivalent.

Definition 6-7 A rule for orienting transverse fiber products of oriented linear spaces is *stable* when, for all such spaces A_1 through A_n and S, for all transverse sequences of maps $(f_1, \ldots, f_n)$ in $\operatorname{Trans}(A_1, \ldots, A_n; S)$, and for any oriented complement C of the fiber product $\vec{P}(f_1, \ldots, f_n) := A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ in the direct product $D := A_1 \times \cdots \times A_n$, there is some neighborhood U of the point $(f_1, \ldots, f_n)$ in $\operatorname{Trans}(A_1, \ldots, A_n; S)$ small enough so that, for all points $(h_1, \ldots, h_n)$ in that neighborhood U, we have

$$\vec{P}(h_1, \ldots, h_n) \oplus C = \vec{P}(f_1, \ldots, f_n) \oplus C. \tag{6-8}$$

Note that the validity of Equation 6-8 combines two claims: first, that the fiber product $\vec{P}(h_1, \ldots, h_n) = A_1[h_1] \times_S \cdots \times_S [h_n]A_n$ is also a complement of the fixed linear subspace C in the direct product $D = A_1 \times \cdots \times A_n$; and second, that using the Concatenate Rule to combine the orientations on $\vec{P}(h_1, \ldots, h_n)$ and C gives the same orientation to the direct product D as does combining the orientations on $\vec{P}(f_1, \ldots, f_n)$ and C.

Exercise 6-9 Using Definition 6-7 to determine what stability means, show that the Greedy-Alice Rule in Figure 1.8 is not stable.

Answer: Let the spaces A, B, and S all be copies of the real numbers $\mathbb{R}$, let $f: A \to S$ be the identity, and let $g: B \to S$ be the zero map. The maps $(f, g) = (1, 0)$ are transverse, and the fiber product $A[1] \times_S [0]B$ is the B-axis of the product space $A \times B$, oriented upward — because Alice is on an upward slope, so Bob should advance. We can take our complement C to be the A-axis, say oriented rightward. But any neighborhood of the point $(f, g) = (1, 0)$ in $\mathrm{Trans}(A, B; S)$ will include points of the form $(h, k) := (1, \epsilon)$, for small ϵ of both signs. When ϵ is positive, the fiber product $A[1] \times_S [\epsilon]B$ has high positive slope and is oriented upward; but, when ϵ is negative, the product $A[1] \times_S [\epsilon]B$ has high negative slope and is oriented downward. In the latter case, the Greedy-Alice Rule does not achieve $(A[1] \times_S [\epsilon]B) \oplus C = (A[1] \times_S [0]B) \oplus C$.

6.3 Lifting from linear spaces to manifolds

Consider some rule for orienting those linear spaces that result as the transverse fiber products of oriented linear spaces. We hope to lift this rule to an analogous rule for fiber products of smooth manifolds. In order for this lifting process to succeed, we have argued that the linear-space rule had better be stable, in the sense of Definition 6-7. In this section, we show that any linear-space rule that is stable in that sense does lift to a smooth-manifold rule.

There is actually a second requirement, in addition to stability; but it almost goes without saying. It is the general mathematical principal that isomorphic inputs should produce isomorphic outputs. In particular, the orientation that our linear-space rule chooses for a transverse fiber product $A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ should depend only upon the dimensions and orientations of the spaces A_1 through A_n and S and the input-output behavior of the factor maps f_1 through f_n. It shouldn't involve flipping a coin. It shouldn't depend upon irrelevant details, such as the colors that the vectors in the various spaces might happen to be colored. We shall formalize this requirement in Section 9.1.1 as the *Isomorphism Axiom*. For now, we argue less formally, referring to an orientation rule that is well-behaved in this sense as *respecting isomorphisms*.

Proposition 6-10 *Any rule for orienting the transverse fiber products of linear spaces that both respects isomorphisms and is stable lifts to a rule for orienting the transverse fiber products of smooth manifolds.*

Proof For notational simplicity, let's focus on the case of binary fiber products. Handling the n-ary case would make the notation more complex without raising any additional mathematical issues.

Fix some linear-space orientation rule that both respects isomorphisms and is stable; and then consider a transverse fiber product of smooth manifolds $\mathbf{P} = \mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$. At any point $\mathbf{p} = (\mathbf{a}, \mathbf{b})$ in $\mathbf{P}$, the tangent space to the fiber product is the fiber product of the tangent spaces:

$$T_{\mathbf{p}}\mathbf{P} \cong (T_{\mathbf{a}}\mathbf{A})[T_{\mathbf{a}}\mathbf{f}] \times_{T_{\mathbf{s}}\mathbf{S}} [T_{\mathbf{b}}\mathbf{g}](T_{\mathbf{b}}\mathbf{B}). \tag{6-11}$$

We orient the tangent space $T_{\mathbf{p}}\mathbf{P}$ on the left by applying our assumed linear-space rule to the fiber product $T_{\mathbf{a}}\mathbf{A} \times_{T_{\mathbf{s}}\mathbf{S}} T_{\mathbf{b}}\mathbf{B}$ on the right. This lifting pseudo-orients the fiber product $\mathbf{P}$; but is this pseudo-orientation locally consistent? We investigate that local consistency in the neighborhood of some particular point, say $\mathbf{p}_0 = (\mathbf{a}_0, \mathbf{b}_0)$ in $\mathbf{P}$. To demonstrate the local consistency near $\mathbf{p}_0$, we construct a chart on some neighborhood of $\mathbf{p}_0$ in $\mathbf{P}$ with which our pseudo-orientation is consistent; and we construct that chart as in Exercise 4-13.

Since we are interested in the structure of the fiber product only near $\mathbf{p}_0$, we can restrict our varying point $\mathbf{p} = (\mathbf{a}, \mathbf{b})$ to vary only within some neighborhood of $\mathbf{p}_0$. We can similarly restrict $\mathbf{a}$ to remain near $\mathbf{a}_0$, $\mathbf{b}$ to remain near $\mathbf{b}_0$, and $\mathbf{s} := \mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{b})$ to remain near $\mathbf{s}_0 := \mathbf{f}(\mathbf{a}_0) = \mathbf{g}(\mathbf{b}_0)$. Choosing charts on those neighborhoods, we might as well assume, as in Exercise 4-13, that the manifold $\mathbf{A}$ actually coincides with an open set in the fixed linear space $\mathbb{R}^a$; we also assume, by translation, that $\mathbf{a}_0 = 0$. We treat the manifolds $\mathbf{B}$ and $\mathbf{S}$ similarly. The fiber product $\mathbf{P}$ is then realized for us as a submanifold of the direct product $\mathbb{R}^a \times \mathbb{R}^b = \mathbb{R}^{a+b}$, and we are interested in what happens on that submanifold near the origin $\mathbf{p}_0 = (0, 0)$. Enforcing these simplifying assumptions formally would involve converting from our initial situation to a different but isomorphic situation. We won't formalize the details of that isomorphism here. Note, however, that our assumed linear-space orientation rule will get the same answer in the simplified situation as in the original situation, since we assume it to respect isomorphisms.

As the point $\mathbf{p}$ varies, the linear spaces $T_{\mathbf{a}}\mathbf{A}$, $T_{\mathbf{b}}\mathbf{B}$, and $T_{\mathbf{s}}\mathbf{S}$ on the right-hand side in Equation 6-11 vary. But now that we are viewing the manifold $\mathbf{A}$ as an open subset of $\mathbb{R}^a$, all of the tangent spaces $T_{\mathbf{a}}\mathbf{A}$ can be naturally identified with the tangent space to $\mathbb{R}^a$, which we shall denote $\tilde{A}$; that is, we have $\tilde{A} = T_{\mathbf{a}}\mathbb{R}^a$ for all points $\mathbf{a}$ near $\mathbf{a}_0 = 0$. Similarly, the tangent spaces $T_{\mathbf{b}}\mathbf{B}$ and $T_{\mathbf{s}}\mathbf{S}$ are all identified with $\tilde{B}$ and $\tilde{S}$, the tangent spaces to $\mathbb{R}^b$ and $\mathbb{R}^s$. With these identifications, we can rewrite Equation 6-11 as

$$T_{\mathbf{p}}\mathbf{P} \cong \tilde{A}[T_{\mathbf{a}}\mathbf{f}] \times_{\tilde{S}} [T_{\mathbf{b}}\mathbf{g}]\tilde{B}.$$

Note that, on this right-hand side, only the factor maps vary.

At the origin $\mathbf{p}_0 = (0, 0)$, we have the tangent space $T_{\mathbf{p}_0}\mathbf{P} \cong \tilde{A}[T_0\mathbf{f}] \times_{\tilde{S}} [T_0\mathbf{g}]\tilde{B}$. Let C denote some oriented complement of this tangent space $T_{\mathbf{p}_0}\mathbf{P}$ in the product space $\tilde{A} \times \tilde{B}$. We saw in Exercise 4-13 that the quotient map from $\mathbb{R}^{a+b}$ to $\mathbb{R}^{a+b}/C$, when restricted to the fiber product $\mathbf{P} = \mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$, serves as a chart on some neighborhood U of the origin $\mathbf{p}_0$ in the fiber product. Let's refer to that chart as $\varphi\colon U \to \mathbb{R}^{a+b}/C$.

As the point $\mathbf{p} = (\mathbf{a}, \mathbf{b})$ varies near $\mathbf{p}_0$, the pair of maps $(T_{\mathbf{a}}\mathbf{f}, T_{\mathbf{b}}\mathbf{g})$ varies away from $(T_0\mathbf{f}, T_0\mathbf{g})$; but the pair $(T_{\mathbf{a}}\mathbf{f}, T_{\mathbf{b}}\mathbf{g})$ remains transverse. Our assumed linear-space orientation rule is stable. Hence, there is some neighborhood of the point $(T_0\mathbf{f}, T_0\mathbf{g})$ in $\mathrm{Trans}(\tilde{A}, \tilde{B}; \tilde{S})$ small enough so that, for every point (h, k) in that neighborhood, we have

$$\big(\tilde{A}[h] \times_{\tilde{S}} [k]\tilde{B}\big) \oplus C = \big(\tilde{A}[T_0\mathbf{f}] \times_{\tilde{S}} [T_0\mathbf{g}]\tilde{B}\big) \oplus C.$$

For all points $\mathbf{p}$ that are close enough to $\mathbf{p}_0$, the pair of maps $(T_{\mathbf{a}}\mathbf{f}, T_{\mathbf{b}}\mathbf{g})$ will belong to that small neighborhood, so we will have

$$\big(\tilde{A}[T_{\mathbf{a}}\mathbf{f}] \times_{\tilde{S}} [T_{\mathbf{b}}\mathbf{g}]\tilde{B}\big) \oplus C = \big(\tilde{A}[T_0\mathbf{f}] \times_{\tilde{S}} [T_0\mathbf{g}]\tilde{B}\big) \oplus C.$$

Thus, if we orient the quotient space $\mathbb{R}^{a+b}/C$ so that our pseudo-orientation is consistent with the chart φ at the origin, that consistency will extend throughout some neighborhood of the origin in the fiber product. $\square$

So the key to an orientation rule for smooth-manifold fiber products is an orientation rule for linear-space fiber products that both respects isomorphisms and is stable. From now on, we focus on the search for a linear-space rule that has those two properties and is otherwise compelling. In particular, we shall work almost entirely with linear spaces and linear maps from now on; smooth manifolds will get mentioned only occasionally.

Chapter 7

The Uncalibrated Delta Rule

Recall that Alice and Bob had two methods for orienting their 1-dimensional fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$: the Partner's-Slope Rule and the Gray-Side Rule.

We generalized the Partner's-Slope Rule to higher dimensions in Section 2.2, getting the Invertible Factor Laws. But those laws are not a complete answer, because there are equidimensional transverse fiber products $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$ about which the Invertible Factor Laws tell us nothing; here is an example.

Exercise 7-1 Consider the torus $\mathbb{S}^1 \times \mathbb{S}^1$, which is a compact smooth manifold. Let $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{S}$ each be a copy of the torus. Viewing each torus as the unit square $[0\ ..\ 1] \times [0\ ..\ 1]$ with edges appropriately identified, define the smooth map $\mathbf{f}\colon \mathbf{A} \to \mathbf{S}$ to be projection on the x axis, $\mathbf{f}(x, y) := (x, 0)$, and define $\mathbf{g}\colon \mathbf{B} \to \mathbf{S}$ to be projection on the y axis, $\mathbf{g}(x, y) := (0, y)$. Show that $\mathbf{f}$ and $\mathbf{g}$ are transverse and describe the fiber product manifold $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$. But note that neither $\mathbf{f}$ nor $\mathbf{g}$ is invertible anywhere.

Answer: At each point in the fiber product $\mathbf{A} \times_{\mathbf{S}} \mathbf{B}$, the tangent spaces and differentials look locally like the example in Exercise 4-5; so the maps $\mathbf{f}$ and $\mathbf{g}$ are transverse, although neither is invertible. The fiber product is another torus, which we can think of as the direct product of the y axis of $\mathbf{A}$ with the x axis of $\mathbf{B}$.

We had better luck when we generalized the Gray-Side Rule in Section 2.3: We sketched out the Delta Rule, a rule that handles every transverse case, regardless of the dimensions of the spaces involved, and that is easily seen to be stable. In this chapter, we flesh out the simplest version of the Delta Rule, the version without an explicit fudge factor. We also define what the Proper Orientation is for any equidimensional transverse fiber product, and we show that our Uncalibrated Delta Rule assigns those Proper Orientations.

7.1 The Delta Rule abstractly

The Gray-Side Rule tells Alice and Bob to walk along the black path so as to keep the gray, Bob-higher side of it to their left. We now generalize that rule to transverse n-ary fiber products in which the dimensions of the various spaces involved are arbitrary. While our overarching goal is an orientation rule for fiber products of smooth manifolds, we focus here on a rule for linear spaces — but we require that rule to be stable, so that it will lift to smooth manifolds.

Consider oriented linear spaces A_1 through A_n and S and transverse linear maps $(f_1, \ldots, f_n)$, with $f_i \colon A_i \to S$. We want to choose an orientation on the fiber product $P := A_1[f_1]\times_S \cdots \times_S [f_n]A_n$, which is a linear subspace of the direct product $D := A_1 \times \cdots \times A_n$. From the given orientations on the factor spaces A_1 through A_n, we can use the Concatenate Rule to orient the direct product D. Let C denote some complement of P in D. If we had a way to orient C, we could combine our orientations on $D = P \oplus C$ and C to get an orientation on P.

The key to orienting C is the difference map $\Delta \colon D \to S^{n-1}$ given by

$$\Delta(\alpha_1, \ldots, \alpha_n) := \bigl(f_2(\alpha_2) - f_1(\alpha_1),\ \ldots,\ f_n(\alpha_n) - f_{n-1}(\alpha_{n-1})\bigr).$$

Recall that the fiber product P is the kernel of the difference map Δ. Furthermore, Definition 4-1 tells us that the maps $(f_1, \ldots, f_n)$ are transverse just when the map Δ is surjective. Hence, the map Δ will carry any complement of $P = \operatorname{Ker}(\Delta)$ in D bijectively onto all of S^{n-1}. Since C is such a complement, we can use the equation $\Delta(C) = S^{n-1}$ to relate an orientation on C to an orientation on S^{n-1}. As for orienting S^{n-1}, we are given a preferred orientation on S and the Concatenate Rule does the rest.

Definition 7-2 (The Uncalibrated Delta Rule) Let n be positive; let S and A_1 through A_n be oriented linear spaces. For i from 1 to n, let $f_i \colon A_i \to S$ be a linear map, and assume that the maps $(f_1, \ldots, f_n)$ are transverse. Let P be the fiber product $P := A_1[f_1]\times_S \cdots \times_S [f_n]A_n$, viewed as a subspace of the direct product $D := A_1 \times \cdots \times A_n$. Let the difference map $\Delta \colon D \to S^{n-1}$ be defined by

$$\Delta(\alpha_1, \ldots, \alpha_n) := \bigl(f_2(\alpha_2) - f_1(\alpha_1), \ldots, f_n(\alpha_n) - f_{n-1}(\alpha_{n-1})\bigr).$$

By transversality, the difference map Δ is surjective. If C denotes any oriented complement of $P = \operatorname{Ker}(\Delta)$ in D, it follows that Δ carries C bijectively onto $\operatorname{Im}(\Delta) = S^{n-1}$. The Uncalibrated Delta Rule orients the fiber product P so that

$$(P \oplus C)/D = \Delta(C)/S^{n-1}. \tag{7-3}$$

In this formula, the denominators $D = A_1 \times \cdots \times A_n$ and S^{n-1} are to be oriented using the Concatenate Rule, and the same for the numerator $P \oplus C$.

Note that which complement C we choose doesn't matter in Equation 7-3, except for its orientation; and even the orientation of C doesn't affect the resulting orientation on the fiber product P, since C appears once in each numerator.

Note also that we have made various arbitrary choices in setting up the Delta Rule. We chose to write the left-hand numerator in Equation 7-3 as $P \oplus C$, rather than $C \oplus P$. In defining the difference map Δ, we chose to subtract $f_i(\alpha_i)$ from $f_{i+1}(\alpha_{i+1})$, rather than the reverse. Because we have made those choices with some care, Equation 7-3, even though it has no fudge factor, turns out to assign the Proper Orientation to P in all equidimensional cases. But getting the Proper Orientations in any-dimensional cases turns out to require explicit calibration.

Exercise 7-4 Consider the unary case $n = 1$. If P is the fiber product of the single space A_1, does the Delta Rule give $P = A_1$ or $P = -A_1$? (The latter would lead to horrible confusion, since our notation $A_1 \times_S \cdots \times_S A_n$ for a fiber product reduces, in the unary case, to A_1.)

Answer: In the unary case, Equation 7-3 reads $(P \oplus C)/A_1 = \Delta(C)/S^0$. The difference map $\Delta\colon A_1 \to S^0$ is the zero map. The empty direct product S^0 is zero-dimensional, and the Concatenate Rule orients it positively; that is, we have $S^0 = \Diamond$. The complement C is zero-dimensional, and we can choose it to be positively oriented as well, giving $C = \Diamond$ and $\Delta(C) = \Diamond$. Equation 7-3 then reduces to $P/A_1 = \Diamond/\Diamond = \Diamond$; thus, the Delta Rule gives $P = A_1$.

7.2 The Delta Rule in matrix form

To ensure that we understand precisely what the Delta Rule means, let's discuss how we would apply it in practice. Which matrices would we construct and which determinants would we compute?

We begin by choosing a basis for each of the linear spaces A_1 through A_n and for S. For simplicity, let's choose those bases to be positively oriented. (Of course, an oriented linear space isomorphic to $-\Diamond$ has no positive bases; but correcting for that glitch is straightforward.) For each i, we then compute the matrix $[f_i]$ of the map f_i, with respect to our chosen bases.

Our next task in implementing the Delta Rule is to assemble these matrices to form the matrix of the map $\Delta\colon A_1 \times \cdots \times A_n \to S^{n-1}$. To get a basis for the domain of Δ, the direct product $A_1 \times \cdots \times A_n$, we do the obvious thing, concatenating our chosen bases for A_1 through A_n, in that order. To get a basis for the codomain

$$S^{n-1} = \underbrace{S \times \cdots \times S}_{n-1},$$

we again do the obvious thing, taking our chosen basis for S and repeating it $n-1$ times, viewed first as a basis for the leftmost S factor, then as a basis for the second S factor, and so forth. Adopting those bases, it is straightforward to check that the matrix of Δ is

$$[\Delta] = \begin{pmatrix} -[f_1] & [f_2] & 0 & \dots & 0 & 0 \\ 0 & -[f_2] & [f_3] & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & [f_{n-1}] & 0 \\ 0 & 0 & 0 & \dots & -[f_{n-1}] & [f_n] \end{pmatrix}$$

The i^{th} column in this formula actually represents a block of a_i columns, while each of the $n-1$ rows represents a block of s rows. So the overall matrix $[\Delta]$ has $d := a_1 + \cdots + a_n$ columns and $(n-1)s$ rows. We next check the rank of the matrix $[\Delta]$. We expect to get $\operatorname{rank}([\Delta]) = (n-1)s$; any lesser rank implies that we do not have transversality.

The fiber product $P := A_1 \times_S \cdots \times_S A_n$ is the kernel of Δ. So our next task is to compute some basis for that kernel. Let $\mathbf{p}$ denote whatever basis we compute; we shall figure out in a moment whether the basis $\mathbf{p}$ is positive or negative, that is, whether $P = \langle \mathbf{p} \rangle$ or $P = -\langle \mathbf{p} \rangle$. The basis $\mathbf{p}$ will consist of $d - \operatorname{rank}([\Delta]) = d - (n-1)s$ vectors in $D = A_1 \times \cdots \times A_n$. Representing each such vector as a column of numbers, the ordered basis $\mathbf{p}$ corresponds to a matrix $[\mathbf{p}]$ with d rows and $d - (n-1)s$ columns. Note that $[\Delta][\mathbf{p}] = 0$.

Next, we compute some complement C of P in D. That is, we compute some matrix $[\mathbf{c}]$ with d rows and $(n-1)s$ columns and with the property that the pasted-together matrix

$$\Big(\Big[\mathbf{p} \Big] \Big[\mathbf{c} \Big] \Big),$$

which is d-by-d in size, has nonzero determinant. The columns of the submatrix $[\mathbf{c}]$ then form a basis for a complement C of P in D. Furthermore, on the left-hand side of Equation 7-3, the quotient $(P \oplus C)/D$ is $\Diamond$ or $-\Diamond$ according as the sign of the determinant of this pasted-together matrix is positive or negative.

To evaluate the quotient $\Delta(C)/S^{n-1}$ on the right-hand side, we form the matrix product $[\Delta][\mathbf{c}]$. This product will be square of size $(n-1)s$, and will have full rank by transversality. The sign of its determinant tells us whether the quotient $\Delta(C)/S^{n-1}$ is $\Diamond$ or $-\Diamond$.

If the signs that we have computed for the left-hand and right-hand sides of Equation 7-3 agree, then our basis $\mathbf{p}$ for P was positive; else it was negative.

Exercise 7-5 Equation 7-3 works when C is chosen to be any complement of P in D. In computational practice, however, it would be attractive to choose C by setting the matrix $[\mathbf{c}]$ to be the transpose of the matrix $[\Delta]$. That is, we choose our complement C to be the linear span of the rows of $[\Delta]$, interpreting each row as a column vector by transposing it. Why is this attractive? What does this choice correspond to geometrically?

Answer: This choice is attractive because the matrix product that arises on the right-hand side is then $[\Delta][\mathbf{c}] = [\Delta][\Delta]^{\mathbf{t}}$. That product is a positive semidefinite matrix and hence has nonnegative determinant automatically; so we know the sign of the determinant without computing it. Geometrically, this strategy corresponds to equipping the space D with the unique Euclidean structure under which our chosen, concatenated basis for D is orthonormal. We then choose C to be the orthogonal complement of P in D — orthogonal under that Euclidean structure.

7.3 Stability of the Delta Rule

A prime virtue of the Delta Rule is its stability, in the sense of Definition 6-7.

Proposition 7-6 *The Uncalibrated Delta Rule is stable. Furthermore, were we to recalibrate the Delta Rule by adding a fudge factor of* ± 1 *that depended only upon the dimensions of the input manifolds, the resulting rule would also be stable.*

Proof Let $P(f_1, \ldots, f_n) := A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ be some transverse fiber product and let C be some oriented complement of the fiber product $P(f_1, \ldots, f_n)$ in the direct product $D := A_1 \times \cdots \times A_n$. The space C will also be a complement of the fiber product $P(h_1, \ldots, h_n) := A_1[h_1] \times_S \cdots \times_S [h_n]A_n$, for every point $(h_1, \ldots, h_n)$ in $\mathrm{Trans}(A_1, \ldots, A_n; S)$ that is close enough to $(f_1, \ldots, f_n)$, say, for those points $(h_1, \ldots, h_n)$ in some neighborhood U of the sequence $(f_1, \ldots, f_n)$. The neighborhood U might not be connected; but the space $\mathrm{Trans}(A_1, \ldots, A_n; S)$ is locally connected, so we will be able to find a smaller neighborhood U' of $(f_1, \ldots, f_n)$ that is included in U and is also connected.

For each point $(h_1, \ldots, h_n)$ in this neighborhood U', the Uncalibrated Delta Rule tells us to orient the fiber product $P(h_1, \ldots, h_n)$ so that

$$\big(P(h_1, \ldots, h_n) \oplus C\big)/D = \Delta_{(h_1,\ldots,h_n)}(C)/S^{n-1}.$$

In particular, we orient $P(f_1, \ldots, f_n)$ so that

$$\big(P(f_1, \ldots, f_n) \oplus C\big)/D = \Delta_{(f_1,\ldots,f_n)}(C)/S^{n-1}.$$

For any point $(h_1, \ldots, h_n)$ in U', the image $\Delta_{(h_1,\ldots,h_n)}(C)$ is all of S^{n-1}, oriented one way or the other. Since the neighborhood U' is connected, we must have either $\Delta_{(h_1,\ldots,h_n)}(C) = S^{n-1}$ for all $(h_1, \ldots, h_n)$ in U' or else $\Delta_{(h_1,\ldots,h_n)}(C) = -S^{n-1}$ for all such $(h_1, \ldots, h_n)$. It follows that $P(h_1, \ldots, h_n) \oplus C = P(f_1, \ldots, f_n) \oplus C$ for all $(h_1, \ldots, h_n)$ in U', which establishes stability.

The same result holds if we recalibrate the Delta Rule by adding a fudge factor of the form $(-1)^{\kappa(a_1,\ldots,a_n;s)}$, where κ is any function $\kappa\colon \mathbb{N}^n \times \mathbb{N} \to \mathbb{Z}/2\mathbb{Z}$. We then orient the fiber product $P(h_1, \ldots, h_n)$ so that

$$\big(P(h_1, \ldots, h_n) \oplus C\big)/D = (-1)^{\kappa(a_1,\ldots,a_n;s)}\, \Delta_{(h_1,\ldots,h_n)}(C)/S^{n-1};$$

but the added fudge factor doesn't vary as the point $(h_1, \ldots, h_n)$ varies, so it doesn't affect stability. □

Because the Delta Rule is stable, Proposition 6-10 says that we can lift it from the category of linear spaces to the category of smooth manifolds — which is what we intend to do.

Corollary 7-7 *Any transverse fiber product* $\mathbf{A}[\mathbf{f}] \times_{\mathbf{S}} [\mathbf{g}]\mathbf{B}$ *of orientable smooth manifolds* $\mathbf{A}$, $\mathbf{B}$, *and* $\mathbf{S}$ *is itself orientable.*

7.4 The Invertible Factor Laws revisited

The Uncalibrated Delta Rule assigns some orientation to every transverse fiber product. But are those orientations the Proper Orientations? As a first stab at that issue, let's analyze whether or not the orientations assigned by the Uncalibrated Delta Rule obey the Invertible Factor Laws. It turns out that they do obey the Right Invertible Law, but they sometimes violate the Left Invertible Law.

7.4.1 The Right Invertible Law

We begin with the Right Invertible Law, which applies only when the dimensions b and s of the linear spaces B and S are equal and the right-hand factor map $g\colon B \to S$ is invertible. In such cases, the map $u\colon A \times_S B \to A$ that projects onto the first coordinate is a bijection, and the Right Invertible Law tell us to orient $A \times_S B$ so that $\operatorname{sgn}(u) = \operatorname{sgn}(g)$, that is, so that $u(A \times_S B)/A = g(B)/S = \operatorname{sgn}(g)\lozenge$. Does the Uncalibrated Delta Rule concur?

When g is invertible, the fiber product $P := A \times_S B$ consists of all pairs of the form $(\alpha, g^{-1}(f(\alpha)))$, for α in A. As a complement C for P in $D = A \times B$, it is convenient to choose all pairs of the form $(0, \beta)$, for β in B. That is, we set $C := (0_{A\leftarrow B}, 1_{B\leftarrow B})(B)$ as an oriented linear space.

We now guess an orientation for P itself, setting $P' := (1_{A\leftarrow A}, g^{-1} \circ f)(A)$. By testing whether or not $(P' \oplus C)/D = \Delta(C)/S$, as in Equation 7-3, we can then determine whether the Uncalibrated Delta Rule sets $P := P'$ or $P := -P'$. For future reference, note that $u(P') = A$, so $u(P')/A = +\lozenge$.

We have $P' \oplus C = \Phi(D) = \Phi(A \times B)$, where $\Phi\colon A \times B \to A \times B$ is the linear map

$$\Phi = \begin{pmatrix} 1_{A\leftarrow A} & 0_{A\leftarrow B} \\ g^{-1} \circ f & 1_{B\leftarrow B} \end{pmatrix},$$

here expressed as a matrix of maps, as discussed in Section 5.4. Note that our recipe for $P' = (1_{A\leftarrow A}, g^{-1} \circ f)(A)$ has become the first column of this 2-by-2 matrix, while our recipe for the complement $C = (0_{A\leftarrow B}, 1_{B\leftarrow B})(B)$ has become its second column. Each column gets formed using Equation 5-6, and the two columns then get pasted together using Equation 5-7. We can easily zero out the entry $g^{-1} \circ f$ in this matrix, either by subtracting an appropriate multiple of the second column from the first or by subtracting an appropriate multiple of the first row from the second. So we have $(P' \oplus C)/D = \operatorname{sgn}(\Phi)\lozenge = +\lozenge$.

What about the quotient $\Delta(C)/S$ on the right-hand side of Equation 7-3? As a matrix of maps, we have

$$\Delta = \begin{pmatrix} -f & g \end{pmatrix}.$$

So we find that

$$\Delta(C) = \begin{pmatrix} -f & g \end{pmatrix} \begin{pmatrix} 0_{A\leftarrow B} \\ 1_{B\leftarrow B} \end{pmatrix} (B) = g(B),$$

which means that $\Delta(C)/S = g(B)/S = \operatorname{sgn}(g)\Diamond$. The Uncalibrated Delta Rule thus insists that $(P \oplus C)/D = \operatorname{sgn}(g)\Diamond$. Since $(P' \oplus C)/D = \Diamond$, we conclude that $P = \operatorname{sgn}(g)P'$ and hence that $u(A \times_S B)/A = u(P)/A = \operatorname{sgn}(g)\, u(P')/A = \operatorname{sgn}(g)\Diamond$, just as the Right Invertible Law requires.

7.4.2 The Left Invertible Law

Symmetrically, the Left Invertible Law applies when $a = s$ and the factor map $f\colon A \to S$ is invertible. In such cases, the projection $v\colon A \times_S B \to B$ onto the second coordinate is a bijection, and the Left Invertible Law tell us to arrange that $\operatorname{sgn}(v) = \operatorname{sgn}(f)$, that is, that $v(A \times_S B)/B = f(A)/S = \operatorname{sgn}(f)\Diamond$. Does the Uncalibrated Delta Rule concur with this? Sadly, not always.

The fiber product $P := A \times_S B$ is the set of pairs $(f^{-1}(g(\beta)), \beta)$, for β in B. We set $P' := (f^{-1} \circ g, 1_{B\leftarrow B})(B)$, noting for future reference that $v(P')/B = \Diamond$.

As our complement C for P in $A \times B$, we take all pairs of the form $(\alpha, 0)$. That is, we set $C := (1_{A\leftarrow A}, 0_{B\leftarrow A})(A)$. We then have $P' \oplus C = \Psi(B \times A)$, where $\Psi\colon B \times A \to A \times B$ is the map

$$\Psi = \begin{pmatrix} f^{-1} \circ g & 1_{A\leftarrow A} \\ 1_{B\leftarrow B} & 0_{B\leftarrow A} \end{pmatrix}.$$

The entry $f^{-1} \circ g$ is easy to zero out; but the map that remains is the swapping isomorphism from $B \times A$ to $A \times B$, which has determinant $(-1)^{ab}$. So we have $(P' \oplus C)/D = (-1)^{ab}\Diamond$.

On the right-hand side, we have

$$\Delta(C) = \begin{pmatrix} -f & g \end{pmatrix} \begin{pmatrix} 1_{A\leftarrow A} \\ 0_{B\leftarrow A} \end{pmatrix} (A) = \neg(f(A)),$$

where $\neg$ is the negating map introduced in Exercise 5-3. Since the dimension of $S = f(A)$ is $s = a$, we end up with $\Delta(C) = (-1)^a f(A)$ and hence we have $\Delta(C)/S = (-1)^a f(A)/S = (-1)^a \operatorname{sgn}(f)$. The Delta Rule thus insists that $(P \oplus C)/D = (-1)^a \operatorname{sgn}(f)\Diamond$. Since $(P' \oplus C)/D = (-1)^{ab}\Diamond$, we conclude that $P = (-1)^{a-ab} \operatorname{sgn}(f)P'$ and hence that

$$v(A \times_S B)/B = v(P)/B = (-1)^{a-ab} \operatorname{sgn}(f)\, v(P')/B = (-1)^{a(1-b)} \operatorname{sgn}(f)\Diamond.$$

When either a is even or b is odd, this concurs with the Left Invertible Law. But when a is odd and b is even, it does not concur. Moral: The Delta Rule must be recalibrated to get the Proper Orientations in the any-dimensional case.

Exercise 7-8 Peeking ahead to Section 9.3, the factor that recalibrates the Delta Rule for the binary, any-dimensional case turns out to be $(-1)^{s(b-s)}$. When $s = a$, verify that this agrees with the $(-1)^{a(1-b)}$ that we just computed.

Hint: An integer, its negation, and its square always have the same parity.

7.5 Equidimensional propriety

The equidimensional case is a happier story: Even the Uncalibrated Delta Rule assigns the Proper Orientation to every equidimensional transverse fiber product. To flesh out this claim, we first have to define what the Proper Orientations are in the equidimensional case, which turns out to be uncontroversial. We then verify that the Uncalibrated Delta Rule produces those Proper Orientations.

7.5.1 The all-invertible case

Before we tackle the full equidimensional case, let's analyze the *all-invertible case*, the subcase in which all of the factor maps are invertible.

Consider an equidimensional fiber product $P := A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ of linear spaces. When all of the factor maps (f_i) are invertible, we shall refer to the fiber product as *all-invertible*. Note that most equidimensional fiber products are all-invertible. If we think of the entries in the matrices of the factor maps as independent variables, those variables have to satisfy a certain algebraic relation in order for $\det([f_i])$ to be zero; so the failure of all-invertibility is already a form of degeneracy. The failure of transversality is even more degenerate, since at least two of the determinants must be zero before the maps $(f_1, \ldots, f_n)$ can become non-transverse, as we saw in Exercise 4-4.

The all-invertible case, in addition to being the most common, is particularly easy to analyze, since we can appeal to the All Invertible Law from Exercise 4-6:

$$\operatorname{sgn}(f_1 \times_S \cdots \times_S f_n) = \operatorname{sgn}(f_1) \cdots \operatorname{sgn}(f_n).$$

Writing our quotients with numerator atop denominator to save space, this law becomes

$$\text{(7-9)} \qquad \frac{(f_1 \times_S \cdots \times_S f_n)(P)}{S} = \frac{f_1(A_1)}{S} \times \cdots \times \frac{f_n(A_n)}{S}.$$

The left-hand side makes sense because, when all of the factor maps are invertible, the fiber-product map $f_1 \times_S \cdots \times_S f_n \colon P \to S$ is also invertible, its inverse taking a vector σ in S to the point $(f_1^{-1}(\sigma), \ldots, f_n^{-1}(\sigma))$ in the fiber product P.

The All Invertible Law tells us that the sign of the fiber-product map should be the product of the signs of the factor maps. That seems elegant and compelling, so we use it to define the Proper Orientation of any fiber product that is all-invertible.

Definition 7-10 When a fiber product $A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ of linear spaces is all-invertible, its *Proper Orientation* is that given by the All Invertible Law 7-9.

Exercise 7-11 Consider an all-invertible fiber product $A_1[f_1] \times_S \cdots \times_S [f_n]A_n$. Show that reversing the orientation of any single factor space A_i reverses the Proper Orientation of the fiber product. But show that reversing the orientation of the base space S reverses the Proper Orientation of the fiber product just when n is even.

Answer: The first claim is clear, since A_i appears just once on the right-hand side of Identity 7-9. For the second claim, note that there are n copies of S on the right-hand side and one on the left-hand side, for a total of $n+1$ copies. When n is even, $n+1$ is odd, so reversing the orientation of S forces us to reverse the orientation of the fiber product. When n is odd, on the other hand, the orientation of the fiber product is independent of the orientation of S.

Note that the second claim continues to hold in the cases $n = 0$ and $n = 1$. The orientation of the nullary fiber product over S, which is S itself, does indeed depend upon the orientation of S, while the orientation of the unary fiber product A_1 over S, which is simply A_1, does not.

We next verify that the Uncalibrated Delta Rule gives the Proper Orientation to every fiber product that is all-invertible.

Theorem 7-12 *Consider any fiber product of linear spaces in which the factor maps are all invertible. The Uncalibrated Delta Rule orients such a fiber product in accord with the All Invertible Law 7-9 and hence gives that fiber product its Proper Orientation.*

Proof Let $P := A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ be a fiber product in which all of the factor maps are invertible. We first use an easy generalization of our analysis of the Right Invertible Law in Section 7.4.1 to show that

$$\frac{u_1(P)}{A_1} = \frac{f_2(A_2)}{S} \times \cdots \times \frac{f_n(A_n)}{S}, \tag{7-13}$$

where the map $u_1\colon P \to A_1$ is the projection from the fiber product to the leftmost factor space A_1. For this portion of the proof, we won't need the assumption that f_1 is invertible. Note that Equation 7-13 is the case $k = 1$ of the All-but-One Invertible Law 4-7, discussed in Exercise 4-6.

Since f_2 through f_n are invertible, the fiber product P consists of all tuples of the form $(\alpha_1,\ f_2^{-1}(f_1(\alpha_1)), \ldots,\ f_n^{-1}(f_1(\alpha_1)))$, for α_1 in A_1. As a complement C for P in $D = A_1 \times \cdots \times A_n$, it is convenient to choose all tuples of the form $(0, \alpha_2, \ldots, \alpha_n)$, for α_i in A_i. More precisely, we set $C := \Phi(A_2 \times \cdots \times A_n)$ as an oriented linear space, where $\Phi\colon A_2 \times \cdots \times A_n \to A_1 \times \cdots \times A_n$ is the obvious injection, the map with matrix

$$[\Phi] = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

We now guess an orientation for the fiber product P itself, setting $P' := (1_{A_1 \leftarrow A_1}, f_2^{-1} \circ f_1, \ldots, f_n^{-1} \circ f_1)(A_1)$. By testing whether or not $(P' \oplus C)/D = \Delta(C)/S^{n-1}$, as in Equation 7-3, we can then determine whether the Uncalibrated Delta Rule sets $P := \pm P'$. For future reference, note that $u_1(P')/A_1 = \Diamond$.

We have $P' \oplus C = \Psi(D) = \Psi(A_1 \times \cdots \times A_n)$, where Ψ is the linear map

$$[\Psi] = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ f_2^{-1} \circ f_1 & 1 & 0 & \ldots & 0 \\ f_3^{-1} \circ f_1 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_n^{-1} \circ f_1 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

obtained by pasting the recipe for P' as a new column at the left of $[\Phi]$. So we have $(P' \oplus C)/D = \operatorname{sgn}(\Psi)\Diamond$. We can easily zero out the subdiagonal entries in the matrix for Ψ; so we have $\operatorname{sgn}(\Psi) = 1$ and hence $(P' \oplus C)/D = \Diamond$.

What about the quotient $\Delta(C)/S^{n-1}$ on the right-hand side of Equation 7-3? The difference map $\Delta\colon D \to S^{n-1}$ has the matrix

$$[\Delta] = \begin{pmatrix} -f_1 & f_2 & 0 & \ldots & 0 & 0 \\ 0 & -f_2 & f_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & f_{n-1} & 0 \\ 0 & 0 & 0 & \ldots & -f_{n-1} & f_n \end{pmatrix},$$

as we saw in Section 7.2. So we have $\Delta(C) = \Delta(\Phi(A_2 \times \cdots \times A_n))$, where

$$[\Delta][\Phi] = \begin{pmatrix} f_2 & 0 & \ldots & 0 & 0 \\ -f_2 & f_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & f_{n-1} & 0 \\ 0 & 0 & \ldots & -f_{n-1} & f_n \end{pmatrix}.$$

Again, we can zero out the subdiagonal entries. So the Uncalibrated Delta Rule insists that

$$\frac{P \oplus C}{D} = \frac{\Delta(C)}{S^{n-1}} = \frac{[\Delta][\Phi](A_2 \times \cdots \times A_n)}{S^{n-1}} = \frac{f_2(A_2)}{S} \times \cdots \times \frac{f_n(A_n)}{S}.$$

Since $u_1(P')/A_1 = (P' \oplus C)/D = \Diamond$, we conclude that

$$\frac{u_1(P)}{A_1} = \frac{u_1(P')}{A_1}\frac{P}{P'} = \frac{P' \oplus C}{D}\frac{P}{P'} = \frac{P \oplus C}{D} = \frac{f_2(A_2)}{S} \times \cdots \times \frac{f_n(A_n)}{S},$$

just as Equation 7-13 requires.

We now suppose that the factor map f_1 is also invertible. We can then rewrite the left-hand quotient $u_1(P)/A_1$ as $f_1(u_1(P))/f_1(A_1)$. Since the fiber-product map $f_1 \times_S \cdots \times_S f_n$ agrees on P with $f_i \circ u_i$ for every i, we can further rewrite this as $(f_1 \times_S \cdots \times_S f_n)(P)/f_1(A_1)$. Multiplying both sides by $f_1(A_1)/S$ then gives us the All-Invertible Law 7-9. $\square$

7.5.2 Extending by continuity

The all-invertible case is now well under control. The All Invertible Law tells us which orientation is the Proper Orientation. And the Delta Rule, even in its current uncalibrated form, computes that Proper Orientation.

On the other hand, the any-dimensional case is still quite an open challenge. In particular, consider a fiber product $A[f] \times_S [g]B$ in which the factor maps f and g are transverse, but neither of them is invertible. The Invertible Factor Laws then don't apply. So it isn't clear how to settle upon one of the two orientations of $A \times_S B$ as the proper one.

Continuity provides a way to settle this question in the equidimensional case. In this section, we show that every transverse, equidimensional fiber product is the limit of all-invertible fiber products. We can define the Proper Orientation of the limit to be the limit of the Proper Orientations, once we have shown that the latter limit is well-defined.

Fix oriented linear spaces A_1 through A_n and S, all of the same dimension, which we shall refer to as s; and consider what happens as we vary the factor maps f_1 through f_n, but keep them always transverse. The point $(f_1, \ldots, f_n)$ then varies in the space $\mathrm{Trans}(A_1, \ldots, A_n; S)$, which we can think of as an open subset of $\mathbb{R}^{a_1 s+\cdots+a_n s} = \mathbb{R}^{ns^2}$. Recall that we can view the fiber-product operator as a map $P\colon \mathrm{Trans}(A_1, \ldots, A_n; S) \to \mathcal{G}(s, A_1 \times \cdots \times A_n)$. If we choose a rule for orienting each fiber product, we get a map $\vec{P}\colon \mathrm{Trans}(A_1, \ldots, A_n; S) \to \vec{\mathcal{G}}(s, A_1 \times \cdots \times A_n)$; and our chosen orientation rule is stable just when this latter map $\vec{P}$ is continuous. Let's settle on the Uncalibrated Delta Rule as our orientation rule, letting $\vec{P}$ denote the continuous map associated with that stable rule.

Let $\mathrm{Inv}(A_1, \ldots, A_n; S)$ denote the subset of $\mathrm{Trans}(A_1, \ldots, A_n; S)$ consisting of those sequences $(f_1, \ldots, f_n)$ in which all of the maps f_1 through f_n are invertible. Claim: The set $\mathrm{Inv}(A_1, \ldots, A_n; S)$ is dense in $\mathrm{Trans}(A_1, \ldots, A_n; S)$.

Exercise 7-14 Let $(f_1, \ldots, f_n)$ be any point in the space $\mathrm{Trans}(A_1, \ldots, A_n; S)$, where we are assuming that the dimensions of A_1 through A_n and S all coincide. Show, by an explicit construction, that the point $(f_1, \ldots, f_n)$ lies on a line segment in the space $\mathrm{Trans}(A_1, \ldots, A_n; S) \subseteq \mathbb{R}^{ns^2}$ all of whose points are all-invertible, with the possible exception of $(f_1, \ldots, f_n)$ itself.

Answer: For each i from 1 to n, there exist lots of linear bijections from A_i to S, since those spaces have the same dimension. Let $h_i\colon A_i \to S$ be one such. We then introduce a real parameter t and we set $f_{i,t} := f_i + t h_i$. This makes $\det([f_{i,t}])$ a polynomial in t with leading term $(\det([h_i]))t^s$; so we will have $\det([f_{i,t}]) = 0$ for at most s different values of t. Thus, all n of the maps $f_{1,t}$ through $f_{n,t}$ will be bijective for all nonzero t with $|t|$ sufficiently small.

The map $\vec{P}$ associated with the Uncalibrated Delta Rule assigns the Proper Orientations to all of the fiber products in $\mathrm{Inv}(A_1, \ldots, A_n; S)$, by Theorem 7-12. Since the map $\vec{P}$ is continuous on the larger set $\mathrm{Trans}(A_1, \ldots, A_n; S)$ in which

$\mathrm{Inv}(A_1, \ldots, A_n; S)$ is dense, it is natural to use continuity in defining the Proper Orientation for this wider class of fiber products.

Definition 7-15 If A_1 through A_n and S are linear spaces of the same dimension and the maps $(f_1, \ldots, f_n)$ are transverse but not all invertible, we define the Proper Orientation of the fiber product $A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ to be the unique orientation that satisfies

$$A_1[f_1] \times_S \cdots \times_S [f_n]A_n = \lim_{t \to 0} A_1[f_{1,t}] \times_S \cdots \times_S [f_{n,t}]A_n, \tag{7-16}$$

where $(f_{1,t}, \ldots, f_{n,t})$ denotes any parameterized sequence of maps that satisfies $\lim_{t\to 0}(f_{1,t}, \ldots, f_{n,t}) = (f_1, \ldots, f_n)$ and that is all-invertible for all nonzero t with $|t|$ sufficiently small.

The continuity of the map $\vec{P}$ implies that the limit on the right-hand side in Equation 7-16 will exist, as an oriented linear space, and that it won't depend upon which parameterized sequence $(f_{1,t}, \ldots, f_{n,t})$ we choose.

Corollary 7-17 *The Uncalibrated Delta Rule assigns the Proper Orientation to every transverse, equidimensional fiber product.*

Note that transversality remains crucial; we can't use continuity to extend even further, to a sequence of maps $(f_1, \ldots, f_n)$ that is not transverse. In such a case, the dimension of the fiber product $A_1[f_1] \times_S \cdots \times_S [f_n]A_n$ is larger than the minimum possible, which is $a_1 + \cdots + a_n - (n-1)s = s$. We may well be able to realize such a sequence $(f_1, \ldots, f_n)$ as the limit of all-invertible sequences $(f_{1,t}, \ldots, f_{n,t})$ in such a way that the limit of the fiber products on the right in Equation 7-16 is well defined. But that limit is then a proper subspace of the fiber product on the left, and hence doesn't help us to orient the latter. Indeed, every s-dimensional subspace of the fiber product on the left, with either orientation, can then be realized as one of the limits on the right.

Chapter 8

Mixed fiber products

The Uncalibrated Delta Rule gives an orientation to every transverse fiber product of linear spaces, whether equidimensional or not. But we were forced to make various arbitrary choices in setting up the Delta Rule, so the answers that it gets aren't always the proper ones. Indeed, we saw in Section 7.4.2 that its answers sometimes violate the Left Invertible Law — a clear sign of impropriety. So we are going to have to recalibrate the Delta Rule, in order to fix up the answers that it gives in the any-dimensional case.

The recalibration itself is easy: We just insert the proper fudge factor. The hard part is deciding what the Proper Orientation of a transverse fiber product ought to be, in general. We hope to find a compelling collection of axioms and a unique orientation rule that satisfies those axioms. If so, we can define the orientations produced by that unique rule to be the Proper Orientations, and we can then adjust the fudge factor in the Delta Rule to produce them.

One obvious axiom to try for is commutativity: $A \times_S B = B \times_S A$. Once we leave the equidimensional case, however, commutativity becomes hopeless. Indeed, we argue in Section 9.1.8 that commutativity is hopeless even for direct products, and direct products are a special case of fiber products: the case in which the base space $S = \lozenge$ is zero-dimensional. The best that we can hope for, when orienting the direct products of oriented linear spaces, is the identity $A \times B = (-1)^{ab} B \times A$. The analogous rule for the more general case of fiber products turns out to be $A \times_S B = (-1)^{(a-s)(b-s)} B \times_S A$.

While we can't require commutativity, we are going to require associativity. In fact, we can actually achieve something stronger than the obvious notion of associativity. The obvious axiom requires $(A \times_S B) \times_S C = A \times_S (B \times_S C)$. We are going to require the stronger identity $(A \times_S B) \times_T C = A \times_S (B \times_T C)$, where the fiber products are *mixed*, meaning that the base spaces S and T may differ. Generalizing to mixed fiber products in this way makes the axiom of associativity significantly more powerful, and that extra power turns out to be just what is needed to narrow down the field of possible orientation rules to a unique, proper choice. In this chapter, we generalize our underlying framework to deal with mixed fiber products.

$$\begin{tikzcd}
A_1 \arrow[dr, "f_1"'] & A_2 \arrow[d, "f_2"] & \cdots & A_n \arrow[dll, "f_n"] \\
& S & &
\end{tikzcd}$$

Figure 8.1: The factor sets and factor maps of a pure fiber product

$$\begin{tikzcd}
A_1 \arrow[dr, "f_1"'] & & A_2 \arrow[dl, "g_2"] \arrow[dr, "f_2"'] & & A_3 \arrow[dl, "g_3"] \arrow[dr, "f_3"'] & & \cdots & & A_n \arrow[dl, "g_n"] \\
& S_1 & & S_2 & & S_3 & \cdots & S_{n-1} &
\end{tikzcd}$$

Figure 8.2: The factor sets and factor maps of a mixed fiber product

8.1 Multiple base spaces

Cast your mind back to Section 4.5 where, in defining the fiber products of sets and set maps, we adopted the particular set $I = \{1, \ldots, n\}$ as our standard index set for an n-ary fiber product. We were then able to switch from writing

$$\prod_{i \in I} {}_S [\,f_i\,] A_i$$

for the fiber product to writing $A_1[\,f_1] \times_S \cdots \times_S [\,f_n] A_n$. And we were able to select, from among all of the fiber-product constraints $f_i(\mathbf{a}_i) = f_j(\mathbf{a}_j)$, the nonredundant constraints $f_i(\mathbf{a}_i) = f_{i+1}(\mathbf{a}_{i+1})$, for i from 1 to $n-1$. Adopting $I = \{1, \ldots, n\}$ as our standard index set also opened the door to a generalization of our underlying framework that we didn't exploit back in Section 4.5: *mixed* fiber products.

Consider Figures 8.1 and 8.2. The fiber products that we have been discussing so far, which we shall henceforth call *pure*, start off with the n factor sets A_1 through A_n and with a single base set S, the i^{th} factor set A_i being equipped with a factor map $f_i \colon A_i \to S$. In contrast, an n-ary *mixed* fiber product starts off with the n factor sets A_1 through A_n and with $n-1$ different base sets, say S_1 through S_{n-1}. The i^{th} factor set A_i comes equipped with two different factor maps, a *forward factor map* $f_i \colon A_i \to S_i$ to the succeeding base set and also a *backward factor map* $g_i \colon A_i \to S_{i-1}$ to the preceding base set. In this way, each of the $n-1$ nonredundant constraints above becomes associated with its own base set. The constraint that we wrote above as $f_i(\mathbf{a}_i) = f_{i+1}(\mathbf{a}_{i+1})$ now becomes $f_i(\mathbf{a}_i) = g_{i+1}(\mathbf{a}_{i+1})$, an equality between elements of the base set S_i. The mixed fiber product

$$A_1[\,f_1] \times_{S_1} [g_2] A_2[\,f_2] \times_{S_2} \cdots \times_{S_{n-1}} [g_n] A_n$$

is that subset of the direct product $A_1 \times \cdots \times A_n$ consisting of those elements $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ for which $f_i(\mathbf{a}_i) = g_{i+1}(\mathbf{a}_{i+1})$, for i from 1 to $n-1$. When the forward and backward fiber maps are clear from the context, we shall often abbreviate a mixed fiber product simply as $A_1 \times_{S_1} \cdots \times_{S_{n-1}} A_n$.

While the first and last factor spaces of a mixed fiber product have only one associated factor map, each of the remaining, interior factor spaces is the domain of two factor maps, one backward and one forward, which we write just before and just after that space. This is in contrast to the pure case, where the interior factor spaces have only one factor map, and we have our choice about whether to write that map just before or just after. (In this monograph, that issue arose only in Exercise 6-5, where we wrote the map g before B in $A[f] \times_S [g]B \times_S [h]C$.)

For binary fiber products, there is no distinction between the pure and mixed cases: Figures 8.1 and 8.2 become indistinguishable when $n = 2$, modulo the choice of names. For unary fiber products, there is technically a distinction, but it makes little difference. In a unary, pure fiber product, the single factor space A_1 is equipped with a factor map $f_1 \colon A_1 \to S$; but the behavior of that map doesn't affect the result. In a unary, mixed fiber product, there are no factor maps.

Most of the theory of fiber products generalizes in a straightforward way from the pure case to the mixed case. We discuss here only those areas where something new happens.

8.2 The fiber-product maps

In the pure case, the fiber-product operation produces, as its outputs, both a set and a map: the set $A_1 \times_S \cdots \times_S A_n$ and the map $f_1 \times_S \cdots \times_S f_n$ from that set to S. In the mixed case, we again get a set: $A_1 \times_{S_1} \cdots \times_{S_{n-1}} A_n$. But there are now $n-1$ different base sets, so there are $n-1$ different fiber-product maps. The i^{th} of these, for i from 1 to $n-1$, takes a point $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ in the fiber-product set to the common value $f_i(\mathbf{a}_i) = g_{i+1}(\mathbf{a}_{i+1})$ in the base set S_i. In the rare situations when we need a name for this map, we'll refer to it as $h_i \colon A_1 \times_{S_1} \cdots \times_{S_{n-1}} A_n \to S_i$.

8.3 The nullary case

In studying pure fiber products, we decided to treat the nullary case specially, but at least it was clear what the nullary pure fiber product over S should be — namely, the space S itself. Once we generalize to mixed fiber products, however, it turns out that there is no natural way to handle the nullary case.

Given a mixed fiber product

$$A_1[f_1] \times_{S_1} [g_2]A_2[f_2] \times_{S_2} \cdots \times_{S_{n-1}} [g_n]A_n,$$

let $g_1 \colon A_1 \to T$ be any map from the factor set A_1 to any nonempty set T, and consider augmenting the product by adding T as an additional factor set and base set on the left, as follows:

$$T[1] \times_T [g_1]A_1[f_1] \times_{S_1} [g_2]A_2[f_2] \times_{S_2} \cdots \times_{S_{n-1}} [g_n]A_n.$$

This augmentation leaves the fiber product essentially unchanged; each point $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ in it is simply augmented to become $(g_1(\mathbf{a}_1), \mathbf{a}_1, \ldots, \mathbf{a}_n)$. Setting $n = 0$ in this augmented product suggests that T might be a natural value for the nullary mixed fiber product. But the set T was arbitrary, except for being nonempty. If the base sets $(S_1, \ldots, S_{n-1})$ came from some natural sequence and that sequence had a zeroth element S_0, then one might argue that the nullary mixed fiber product should be S_0. But it would make equal sense to augment from the right, which would argue for S_n instead. And, when the base sets (S_i) are arbitrary, there is no reason to prefer any particular set for T. Therefore, when dealing with mixed fiber products, we flatly outlaw the nullary case.

In almost all respects, mixed fiber products are more general than pure fiber products. Indeed, we can restrict ourselves from the mixed case to the pure case by constraining all of the base spaces to be equal: $S_1 = \cdots = S_{n-1}$; and by constraining the forward and backward maps defined on any common factor space to coincide: $f_i = g_i$ for $1 < i < n$. The nullary case is the lone exception. When we restrict ourselves to pure fiber products over S, we are limiting ourselves to a world in which S is the only permissible base space; so S then becomes a natural value for the nullary fiber product. It is precisely the fact that the mixed case is more general, allowing lots of different base spaces, that forces us to abandon the notion of a nullary fiber product.

8.4 Commutativity and associativity

In the pure case, the operator that takes fiber products over a fixed base space S is both commutative and associative. In the mixed case, on the other hand, each fiber product involves a sequence of base spaces, each of which is related to the factor spaces on either side of it by forward and backward factor maps. So there is no hope for commutativity in general; the left-to-right order of the factor spaces and of the base spaces is encoded in the factor maps.

What about associativity, in the mixed case? It takes a moment's thought to see that mixed fiber products can even be nested. Given indices i and j with $1 \le i \le j \le n$ and given the n-ary mixed product

$$A_1[\,f_1] \times_{S_1} \cdots \times_{S_{i-1}} [g_i]A_i[\,f_i] \times_{S_i} \cdots \times_{S_{j-1}} [g_j]A_j[\,f_j] \times_{S_j} \cdots \times_{S_{n-1}} [g_n]A_n,$$

suppose that we want to compute the product from A_i through A_j first, as an inner fiber product $P := A_i[\,f_i] \times_{S_i} \cdots \times_{S_{j-1}} [g_j]A_j$. Roughly speaking, this corresponds to adding in a pair of parentheses, getting

$$A_1[\,f_1] \times_{S_1} \cdots \times_{S_{i-1}} [g_i'] \big(A_i[\,f_i] \times_{S_i} \cdots \times_{S_{j-1}} [g_j]A_j\big)[\,f_j'] \times_{S_j} \cdots \times_{S_{n-1}} [g_n]A_n,$$

or, equivalently,

$$A_1[\,f_1] \times_{S_1} \cdots \times_{S_{i-1}} [g_i']\, P\, [\,f_j'] \times_{S_j} \cdots \times_{S_{n-1}} [g_n]A_n.$$

For this remaining, outer product to make sense, however, the set P must come equipped with a backward factor map to S_{i-1} and a forward factor map to S_j, the maps that we have written above as g'_i and f'_j. To define those maps, we exploit the fact that the inner fiber product P is a subset of the direct product $D := A_i \times \cdots \times A_j$. Let $u_i \colon D \to A_i$ and $u_j \colon D \to A_j$ be the projections of D onto its first and last components. We then define the backward factor map $g'_i \colon P \to S_{i-1}$ by setting $g'_i := g_i \circ u_i$ and the forward factor map $f'_j \colon P \to S_j$ by setting $f'_j := f_j \circ u_j$. With this understanding, it is easy to check that mixed fiber products can be nested and that the mixed-fiber-product operator is associative.

8.5 Linear spaces

We now specialize to the category of linear spaces and linear maps. The input data for a mixed fiber product in that category is two sequences of linear spaces, connected by maps in the pattern of Figure 8.2. We shall refer to such a structure as a *zigzag*. When we need to write a zigzag down in text, it saves space to collapse all of its spaces and maps down onto a single line, like this:

$$A_1 \xrightarrow{f_1} S_1 \xleftarrow{g_2} A_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} S_{n-1} \xleftarrow{g_n} A_n.$$

We could allow zigzags to end either on the top or on the bottom, that is, to end either at a factor space A_i or at a base space S_i. But it is easy to pad a zigzag, at either end, by adjoining zero-dimensional linear spaces and identically zero linear maps; so it doesn't matter too much where the zigzag officially ends. Since we here use zigzags mostly as input to the mixed-fiber-product operator, we'll require our zigzags to have both their left and right ends on top. We refer to $A_1 \overset{f_1}{\to} S_1 \overset{g_2}{\leftarrow} A_2 \overset{f_2}{\to} \cdots \overset{f_{n-1}}{\to} S_{n-1} \overset{g_n}{\leftarrow} A_n$ as an *n-ary zigzag*, the input data for an n-ary fiber product. So a unary zigzag consists of a single factor space A_1, with no base spaces and no factor maps. (And there is no such thing as a nullary zigzag.)

Just as for pure fiber products, the linearity of the i^{th} base space S_i in a mixed, n-ary fiber product lets us do a subtraction to measure the extent to which the i^{th} of the nonredundant constraints fails to hold. We define the *difference map* for the zigzag $A_1 \overset{f_1}{\to} S_1 \overset{g_2}{\leftarrow} A_2 \overset{f_2}{\to} \cdots \overset{f_{n-1}}{\to} S_{n-1} \overset{g_n}{\leftarrow} A_n$ to be the map $\Delta \colon A_1 \times \cdots \times A_n \to S_1 \times \cdots \times S_{n-1}$ given by

$$\Delta(\alpha_1, \ldots, \alpha_n) := \big(g_2(\alpha_2) - f_1(\alpha_1),\ \ldots,\ g_n(\alpha_n) - f_{n-1}(\alpha_{n-1})\big).$$

As in the pure case, we define a zigzag to be *transverse* just when the associated difference map is surjective. It is precisely in the transverse case that the mixed fiber product $A_1 \times_{S_1} A_2 \times_{S_2} \cdots \times_{S_{n-1}} A_n = \operatorname{Ker}(\Delta)$ has the minimum possible dimension, which is $(a_1 + \cdots + a_n) - (s_1 + \cdots + s_{n-1})$.

Note that generalizing from pure to mixed simplifies some parts of the theory. The difference map in the mixed case takes its values in the space $S_1 \times \cdots \times S_{n-1}$, which is the direct product of all of the base spaces. In the pure case, because $S_1 = \cdots = S_{n-1} = S$, that product collapses to S^{n-1}, which is more mysterious. Similarly, the formula $(a_1 + \cdots + a_n) - (s_1 + \cdots + s_{n-1})$ for the dimension of a transverse, mixed fiber product is more enlightening than the corresponding formula $(a_1 + \cdots + a_n) - (n-1)s$ in the pure case.

In the pure case, we can think of transversality as a property of a sequence of S-valued linear maps. In the mixed case, however, if we want to think of transversality as a property about maps, we have to start with a sequence of pairs of maps, the two maps in each pair being the forward and backward factor maps whose values lie in a common base space. For example, the transversality of the n-ary zigzag above, if we think in terms of maps, is the transversality of the maps $((f_1, g_2), (f_2, g_3), \ldots, (f_{n-1}, g_n))$.

Proposition 4-9 showed that the property of transversality is associative, in the pure case. That associativity carries over to the mixed case, and for the same reason: Suppose that we compute an overall fiber product in a nested fashion, first combining some of the factor spaces in an inner fiber product and then combining that result with the remaining factor spaces in an outer fiber product. The overall product is transverse just when both the inner and outer products are transverse, since both of those events happen just when the final fiber-product space has the minimum possible dimension.

To ensure that this is clear, let's state the mixed-case analog of Proposition 4-9. The proof for the mixed case is essentially the same as for the pure case.

Proposition 8-1 *Let $A_1 \xrightarrow{f_1} S_1 \xleftarrow{g_2} A_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} S_{n-1} \xleftarrow{g_n} A_n$ be an n-ary zigzag, and let i and j be integers with $1 \le i \le j \le n$. Let P denote the inner, mixed fiber product $P := A_i[f_i] \times_{S_i} [g_{i+1}]A_{i+1}[f_{i+1}] \times_{S_{i+1}} \cdots \times_{S_{j-1}} [g_j]A_j$ of the factor spaces from A_i through A_j. The overall, n-ary zigzag is transverse just when*

- *the zigzag $A_i \xrightarrow{f_i} S_i \xleftarrow{g_{i+1}} A_{i+1} \xrightarrow{f_{i+1}} \cdots \xrightarrow{f_{j-1}} S_{j-1} \xleftarrow{g_j} A_j$ associated with inner fiber product P is transverse*

- *and the zigzag $A_1 \xrightarrow{f_1} \cdots \xrightarrow{f_{i-1}} S_{i-1} \xleftarrow{g_i'} P \xrightarrow{f_j'} S_j \xleftarrow{g_{j+1}} \cdots \xleftarrow{g_n} A_n$ associated with the outer fiber product is transverse. The space P is a single factor space of this outer fiber product, and its forward and backward factor maps are the maps f_j' and g_i' defined in Section 8.4.*

Exercise 8-2 Consider a ternary zigzag $A \xrightarrow{f} S \xleftarrow{g} B \xrightarrow{h} T \xleftarrow{k} C$. If we compute the mixed fiber product $A[f] \times_S [g]B[h] \times_T [k]C$ in one step, the associated difference map will be the map $\Delta\colon A \times B \times C \to S \times T$ given by $\Delta(\alpha, \beta, \gamma) := (g(\beta) - f(\alpha), k(\gamma) - h(\beta))$. If we compute that same product in the nested fashion $\big(A[f] \times_S [g]B\big)[h'] \times_T [k]C$, the inner and outer fiber products have the difference maps $\Delta_I\colon A \times B \to S$ and $\Delta_O\colon (A[f] \times_S [g]B) \times C \to T$ given by

$$\Delta_I(\alpha, \beta) := g(\beta) - f(\alpha)$$
$$\Delta_O((\alpha, \beta), \gamma) := k(\gamma) - h'(\alpha, \beta) = k(\gamma) - h(\beta).$$

Proposition 8-1 tells us that the overall difference map Δ will be surjective just when both the inner and outer difference maps Δ_I and Δ_O are surjective. Verify this directly.

Answer: Suppose first that Δ is surjective. To see that Δ_I is surjective, for any σ in S, chose an arbitrary τ in T; there exists (α, β, γ) with $\Delta(\alpha, \beta, \gamma) = (\sigma, \tau)$, from which it follows that $\Delta_I(\alpha, \beta) = \sigma$. To see that Δ_O is surjective, for any τ in T, there exists (α, β, γ) with $\Delta(\alpha, \beta, \gamma) = (0, \tau)$. From this, it follows that the pair (α, β) lies in the fiber product $A[f] \times_S [g]B$ and that $\Delta_O((\alpha, \beta), \gamma) = \tau$.

Conversely, suppose that both Δ_I and Δ_O are surjective, and let (σ, τ) be any point in $S \times T$. There exists (α_1, β_1) with $\Delta_I(\alpha_1, \beta_1) = \sigma$. And then there exists $((\alpha_2, \beta_2), \gamma_2)$ with $f(\alpha_2) = g(\beta_2)$ and $\Delta_O((\alpha_2, \beta_2), \gamma_2) = \tau + h(\beta_1)$. From these, it follows that $\Delta(\alpha_1 + \alpha_2, \beta_1 + \beta_2, \gamma_2) = (\sigma, \tau)$.

Exercise 8-3 Generalize the Invertible Factor Laws from the n-ary pure case, as in Exercise 4-6, to the n-ary mixed case.

Answer: In the mixed fiber product

$$P := A_1[f_1] \times_{S_1} [g_2]A_2[f_2] \times_{S_2} \cdots \times_{S_{n-1}} [g_n]A_n,$$

let $h_i \colon P \to S_i$ denote the i^{th} fiber-product map, for i from 1 to $n - 1$. And let $u_i \colon P \to A_i$ denote the i^{th} projection, for i from 1 to n. So we have $h_i = f_i \circ u_i = g_{i+1} \circ u_{i+1}$, for i from 1 to $n - 1$. And we then have

$$\begin{aligned}
\operatorname{sgn}(u_i) &= \operatorname{sgn}(f_1) \cdots \operatorname{sgn}(f_{i-1}) \operatorname{sgn}(g_{i+1}) \cdots \operatorname{sgn}(g_n) \\
\operatorname{sgn}(u_{i+1}) &= \operatorname{sgn}(f_1) \cdots \operatorname{sgn}(f_i) \operatorname{sgn}(g_{i+2}) \cdots \operatorname{sgn}(g_n) \\
\operatorname{sgn}(h_i) &= \operatorname{sgn}(f_1) \cdots \operatorname{sgn}(f_i) \operatorname{sgn}(g_{i+1}) \cdots \operatorname{sgn}(g_n).
\end{aligned}$$

In any of these formulas, when all of the maps on the right-hand side are invertible, then the map on the left will also be invertible, and we should orient the mixed fiber product P so as to make the equality hold.

8.6 Stable orientation rules

If we fix the factor spaces A_1 through A_n and the base spaces S_1 through S_{n-1}, what is still needed to specify a mixed fiber product are the forward factor maps $(f_1, \ldots, f_{n-1})$ and the backward factor maps $(g_2, \ldots, g_n)$. Let's define

$$\operatorname{Trans}(A_1, \ldots, A_n; S_1, \ldots, S_{n-1})$$

to be those sequences of maps $(f_1, \ldots, f_{n-1}; g_2, \ldots, g_n)$ that are transverse, that is, those for which the resulting difference map Δ is surjective. Letting $p := (a_1 + \cdots + a_n) - (s_1 + \cdots + s_{n-1})$ denote the dimension of the resulting fiber products, we can think of the mixed-fiber-product operator as a mapping

$$P \colon \operatorname{Trans}(A_1, \ldots, A_n; S_1, \ldots, S_{n-1}) \to \mathcal{G}(p, A_1 \times \cdots \times A_n).$$

If the spaces A_1 through A_n and S_1 through S_{n-1} are oriented, then an orientation rule for mixed fiber products converts this into a mapping

$$\vec{P}\colon \operatorname{Trans}(A_1, \ldots, A_n; S_1, \ldots, S_{n-1}) \to \vec{\mathcal{G}}(p, A_1 \times \cdots \times A_n).$$

An orientation rule is *stable* just when this mapping is continuous.

8.7 Smooth manifolds

Once we understand mixed fiber products of linear spaces, we could go on to study mixed fiber products of smooth manifolds. In the transverse case, such a fiber product will again be a smooth manifold, by an easy generalization of Proposition 4-12. (Indeed, if we evaluate an n-ary mixed fiber product as a nested sequence of $n-1$ binary fiber products, we can then apply Proposition 4-12 as stated, since, for binary products, there is no distinction between pure and mixed.) Furthermore, by an easy generalization of Proposition 6-10, any orientation rule for the transverse, mixed fiber products of linear spaces that is stable and that respects isomorphisms will lift to an orientation rule for the transverse, mixed fiber products of smooth manifolds. But it wouldn't require any new ideas to push the theory in this direction, so we shan't bother to do so. Instead, we shall continue to focus on linear spaces and linear maps.

Chapter 9

Propriety via axioms

So our new goal is as follows: Given any transverse zigzag $A_1 \xrightarrow{f_1} S_1 \xleftarrow{g_2} A_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} S_{n-1} \xleftarrow{g_n} A_n$ in which all of the factor spaces A_1 through A_n and base spaces S_1 through S_{n-1} are oriented, find a rule that orients the mixed fiber product $A_1 \times_{S_1} A_2 \times_{S_2} \cdots \times_{S_{n-1}} A_n$. We want our rule to satisfy as many pretty axioms as possible. In particular, we require that the rule be stable and that it respect isomorphisms, so that it will lift to give us an analogous rule for orienting the transverse, mixed fiber products of smooth manifolds.

The subtle issue is finding a family of axioms that is consistent and complete. Consistent means that there is some orientation rule that satisfies all of our axioms. Complete means that there is only one such rule; that is, we have imposed enough axioms to eliminate any flexibility in the choice of the orientation rule. Once we have a consistent and complete family of axioms, it is straightforward to construct the unique orientation rule that satisfies them: We simply calibrate the Delta Rule (after generalizing it to the mixed case) so that it satisfies our axioms.

By the way, we shall feel free to add in lots of axioms, provided that they are consistent, in our quest for completeness. In particular, we make no attempt to arrange that each of our axioms be independent of the others.

9.1 The axioms

In stating our axioms, it is convenient to focus on the binary case. Define a *binary problem instance* to be a transverse, binary zigzag $A \xrightarrow{f} S \xleftarrow{g} B$ whose spaces A, B, and S are oriented. A *binary orientation rule* chooses, for every binary problem instance $A \xrightarrow{f} S \xleftarrow{g} B$, one of the two possible orientations on the fiber product $A[f] \times_S [g]B$. We give our axioms as restrictions on the behavior of a binary orientation rule.

Define a *ternary problem instance*, analogously, to be a transverse, ternary zigzag $A \xrightarrow{f} S \xleftarrow{g} B \xrightarrow{h} T \xleftarrow{k} C$ whose five spaces are oriented. The associativity law $(A \times_S B) \times_T C = A \times_S (B \times_T C)$ for ternary problem instances will be one of our axioms. Because of this, there is no need for us to define a notion of a "ternary orientation rule", or an "n-ary orientation rule" for any $n > 2$. If a binary orientation rule satisfies that associativity axiom, then we can extend it uniquely into an orientation rule for transverse, mixed fiber products $A_1 \times_{S_1} A_2 \times_{S_2} \cdots \times_{S_{n-1}} A_n$ of any arity n, since all ways of evaluating such an n-ary fiber product using $n-1$ nested binary fiber products will give the same result.

9.1.1 The Isomorphism Axiom

We want our rule for orienting the fiber product $A[f] \times_S [g]B$ to be based solely on the abstract structure of the linear spaces A, B, and S, the orientations on those linear spaces, and the abstract structure of the linear maps f and g. We don't want the rule to flip a coin. We don't want its decision influenced by extraneous properties, such as the colors that the elements of A, B, or S might happen to be painted. To enforce these restrictions, we insist that the rule *respect isomorphisms*, that is, that it give isomorphic answers to isomorphic problem instances.

If $A \xrightarrow{f} S \xleftarrow{g} B$ and $A' \xrightarrow{f'} S' \xleftarrow{g'} B'$ are two binary problem instances, what does it mean for them to be isomorphic? No surprise: They are isomorphic just when there exist orientation-preserving bijections between corresponding spaces that commute with the factor maps:

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "\varphi"'] & S \arrow[d, "\chi"'] & B \arrow[l, "g"'] \arrow[d, "\psi"'] \\ A' \arrow[r, "f'"] & S' & B' \arrow[l, "g'"'] \end{tikzcd} \tag{9-1}$$

That is, there must exist maps $\varphi \colon A \to A'$, $\psi \colon B \to B'$, and $\chi : S \to S'$ that are orientation-preserving bijections and that satisfy $\chi \circ f = f' \circ \varphi$ and $\chi \circ g = g' \circ \psi$.

Given two problem instances $A \xrightarrow{f} S \xleftarrow{g} B$ and $A' \xrightarrow{f'} S' \xleftarrow{g'} B'$ that are isomorphic, the resulting fiber products are also isomorphic, at least up to orientation. In particular, consider the map $\omega \colon (A \times_S B) \to (A' \times_{S'} B')$ defined by $\omega(\alpha, \beta) := (\varphi(\alpha), \psi(\beta))$. This definition makes sense because the fiber product $A \times_S B$ is a subset of the direct product $A \times B$ and because the unprimed constraint $f(\alpha) = g(\beta)$ implies the primed constraint $f'(\varphi(\alpha)) = \chi(f(\alpha)) = \chi(g(\beta)) = g'(\psi(\beta))$. The map ω is easily seen to be bijective. A binary orientation rule satisfies the *Isomorphism Axiom* just when, in this situation, it always orients the unprimed and primed fiber products so that ω preserves orientation.

9.1.2 The Stability Axiom

We want our rule for orienting linear-space fiber products to lift to a rule for orienting smooth-manifold fiber products. Hence, as we discussed in Chapter 6, we must require stability. We enshrine this requirement as the *Stability Axiom*: Our orientation rule for binary problem instances must be *stable*, in the sense of the equivalent Definitions 6-6 and 6-7.

9.1.3 The Reversing Axioms

Reversing the orientation on any one of the spaces A, B, or S, while leaving the factor maps f and g unchanged, must reverse the orientation on the fiber product $A[f] \times_S [g]B$. In symbols, we can write this *Reversing Axiom* as follows:

$$(-A) \times_S B = A \times_S (-B) = A \times_{(-S)} B = -(A \times_S B).$$

We shall sometimes split this axiom into three pieces: the *Left Reversing Axiom*, the *Right Reversing Axiom*, and the *Base Reversing Axiom*.

Reversing the orientation of a base space is another situation where mixed fiber products are simpler than pure ones. Note that a mixed, n-ary fiber product $A_1 \times_{S_1} A_2 \times_{S_2} \cdots \times_{S_{n-1}} A_n$ has $n-1$ different base spaces, S_1 through S_{n-1}. Reversing any one of them reverses the fiber product — no mystery there. In the pure case, however, our only option is to reverse all $n-1$ of them simultaneously, since they are all copies of a single base space S. That explains why, as we noted in Exercise 7-11, the orientation of a pure, n-ary fiber product $A_1 \times_S \cdots \times_S A_n$ depends on the orientation of the base space S just when $n-1$ is odd.

9.1.4 The Both Identities Axiom

Consider the particular binary problem instance $S \xrightarrow{1} S \xleftarrow{1} S$, in which both of the factor spaces A and B coincide with the base space S and both of the factor maps f and g are the identity $1\colon S \to S$. The fiber product $S \times_S S$ is naturally isomorphic to S, and the fiber-product map $1 \times_S 1$ is that isomorphism: the map that takes (σ, σ) to σ. The *Both Identities Axiom* requires that the space $S \times_S S$ be oriented so that the map $1 \times_S 1$ preserves orientation. In symbols, we shall write this axiom simply as

$$S \times_S S = S,$$

where each factor map is understood to be the identity.

If we combine the Both Identities Axiom with the Left and Right Reversing Axioms, we deduce that $(-S) \times_S S = S \times_S (-S) = -S$, while $(-S) \times_S (-S) = S$. If we further combine those axioms with the Isomorphism Axiom, we can handle any problem instance $A \xrightarrow{f} S \xleftarrow{g} B$ in which both f and g are invertible; any such problem instance is isomorphic to one of the four particular instances

$$\begin{aligned}
S &\xrightarrow{1} S \xleftarrow{1} S\\
S &\xrightarrow{1} S \xleftarrow{1} (-S)\\
(-S) &\xrightarrow{1} S \xleftarrow{1} S\\
(-S) &\xrightarrow{1} S \xleftarrow{1} (-S),
\end{aligned}$$

according as the maps f and g either preserve or reverse orientation.

Exercise 9-2 One reason why the equidimensional case is easy is that we have already listed enough axioms to determine that case completely. Suppose that we require the Isomorphism Axiom, the Stability Axiom, the Left and Right Reversing Axioms, and the Both Identities Axiom. Show that every equidimensional, binary, transverse fiber product then has a well-determined orientation.

Sketch: By the Stability Axiom, it suffices to consider problem instances in which both of the factor maps f and g are invertible. By the Left and Right Reversing Axioms, we can further assume that both f and g preserve orientation. By the Isomorphism Axiom, we can then assume that we have the particular problem instance $S \xrightarrow{1} S \xleftarrow{1} S$, whose answer is $+S$ by the Both Identities Axiom.

Exercise 9-3 In equidimensional cases, show that the *Commutativity Axiom*, the axiom $A \times_S B = B \times_S A$, follows from our other axioms.

Answer: By the Stability and Isomorphism Axioms, it suffices to consider the four cases in which A and B are either $+S$ or $-S$, with both factor maps f and g being the identity $1\colon S \to S$. By the Reversing and Both Identities Axioms, those four cases do commute.

9.1.5 The Left Identity Axiom

In order to nail down the Proper Orientations for problem instances that are not equidimensional, we need further axioms. One thing to consider is what should happen when one of the factor maps is the identity, but the other is not.

Consider a binary problem instance of the form $S \xrightarrow{1} S \xleftarrow{g} B$, in which the left factor space A coincides with the base S and the left factor map f is the identity $1\colon S \to S$, but the right factor space and factor map are unconstrained. In such a case, projection onto the right factor is an isomorphism between the fiber product $S \times_S B$ and the right factor space B, and the *Left Identity Axiom* requires that $S \times_S B$ be oriented so as to make this isomorphism preserve orientation. In symbols, we encode this axiom as the formula $S \times_S B = B$. From another perspective, this formula asserts that the identity map $1\colon S \to S$ should act as a left identity for the fiber product, viewed as a binary operation on maps to S.

The Left Identity Axiom combines with the Left Reversing Axiom to give us the formula $(-S) \times_S B = -B$. If we add in the Isomorphism Axiom, we can handle any problem instance in which the left factor map f is invertible, whether it preserves or reverses orientation. Thus, we now have sufficient axioms to guarantee that our orientation rule will obey the Left Invertible Law.

The Left Identity Axiom is stronger than the Both Identities Axiom, of course, the latter being the special case of the former in which $B := S$ and $g := 1$.

9.1.6 The Right Identity Axiom

The *Right Identity Axiom* is symmetric. It requires that $A \times_S S = A$, and imposing it guarantees that our orientation rule will obey the Right Invertible Law.

9.1.7 The Axiom of Mixed Associativity

The key axiom that nails down the Proper Orientations in non-equidimensional cases is the Axiom of Mixed Associativity. Here is where we reap our reward for generalizing from pure to mixed fiber products.

Let $A \xrightarrow{f} S \xleftarrow{g} B \xrightarrow{h} T \xleftarrow{k} C$ be a ternary problem instance. Since the maps $((f, g), (h, k))$ are then transverse, it follows, as in Exercise 8-2, that all four of the following are binary problem instances with transverse pairs of maps:

$$A \xrightarrow{f} S \xleftarrow{g} B$$

$$(A \times_S B) \xrightarrow{h'} T \xleftarrow{k} C$$

$$B \xrightarrow{h} T \xleftarrow{k} C$$

$$A \xrightarrow{f} S \xleftarrow{g'} (B \times_T C)$$

Here, we have written h' for the forward factor map on the second line just to remind ourselves that we must project from the fiber product $A \times_S B$ onto its right-hand factor B before we can apply the map h to get to T. The map g' on the fourth line is similar. Since these four binary problems are transverse, the two nested fiber products $(A \times_S B) \times_T C$ and $A \times_S (B \times_T C)$ are well defined. They are also canonically isomorphic. The *Axiom of Mixed Associativity* requires that they be oriented so that the canonical isomorphism preserves orientation. In symbols, we have $(A \times_S B) \times_T C = A \times_S (B \times_T C)$.

Warning: Associativity is often used as a rewrite rule; but this axiom has a side condition that we must check, before doing such a rewriting. Suppose that we come across a nested, binary fiber product in some calculation of ours, say $(A \times_S B) \times_T C$. Before we can use the Axiom of Mixed Associativity to move the parentheses to the right, we must check that we actually have a ternary problem instance. Transversality isn't typically a problem, since, when the inner and outer binary products are transverse, the ternary product is also transverse. But there is a more basic reason why there might not be a ternary product, an issue about the left factor map of the outer product. Let's call that map $e\colon A \times_S B \to T$, so our current product is actually $(A \times_S B)[e] \times_T C$. To apply the Axiom of Mixed Associativity, we need a factor map $h\colon B \to T$. Our context may well supply such a map h, a map with the property that $e(\mathbf{a}, \mathbf{b}) = h(\mathbf{b})$, for every point $(\mathbf{a}, \mathbf{b})$ in $A \times_S B$. Our map e can then play the role of h' and the Axiom of Mixed Associativity will apply. But we must check for the existence of the map h as a side condition. If the value $e(\mathbf{a}, \mathbf{b})$ depended on $\mathbf{a}$, as well as on $\mathbf{b}$, then no such map h could be defined and the Axiom of Mixed Associativity wouldn't apply.

9.1.8 The issue of commutativity

Commutativity is a frequent companion to associativity. But there is no hope for an oriented fiber-product operation to be commutative in full generality.

Proposition 9-4 *A rule for orienting the transverse fiber products of linear spaces, if it satisfies the Isomorphism Axiom, cannot be fully commutative.*

Proof In the hoped-for identity $A \times_S B = B \times_S A$, consider a case in which $S := \lozenge$ is zero-dimensional and positively oriented, while both A and B are 1-dimensional, say with $A = \langle\alpha\rangle$ and $B = \langle\beta\rangle$. The two factor maps f and g must be identically zero, and those maps are transverse. The fiber-product $A \times_S B$ is the entire plane $A \times B$, for which the vectors $(\alpha, 0)$ and $(0, \beta)$ form an obvious basis. Our orientation rule will determine which comes first, $(\alpha, 0)$ or $(0, \beta)$, in a positive basis for the plane $A \times_S B$. Letting η denote ± 1 as appropriate, we have

$$A \times_S B = \eta\langle(\alpha, 0), (0, \beta)\rangle. \tag{9-5}$$

What would commutativity mean, in this case? Let $\Upsilon\colon A \times B \to B \times A$ be the swapping map, the map defined by $\Upsilon(x\alpha, y\beta) := (y\beta, x\alpha)$, for all real numbers x and y. For commutativity to hold, we would have to have $\Upsilon(A \times_\lozenge B) = B \times_\lozenge A$. Substituting in from Equation 9-5, we would need

$$B \times_\lozenge A = \Upsilon\big(\eta\langle(\alpha, 0), (0, \beta)\rangle\big) = \eta\langle(0, \alpha), (\beta, 0)\rangle. \tag{9-6}$$

But our original problem instance $A \stackrel{0}{\to} \lozenge \stackrel{0}{\leftarrow} B$ is also isomorphic to the swapped problem instance $B \stackrel{0}{\to} \lozenge \stackrel{0}{\leftarrow} A$ as follows:

$$\begin{tikzcd} A \arrow[r, "0"] \arrow[d, "\varphi"'] & \lozenge \arrow[d, "0"'] & B \arrow[l, "0"'] \arrow[d, "\psi"'] \\ B \arrow[r, "0"] & \lozenge & A \arrow[l, "0"'] \end{tikzcd}$$

Here, the map $\varphi\colon A \to B$ is given by $\varphi(t\alpha) := t\beta$, for all real numbers t; and the map $\psi\colon B \to A$ is $\psi := \varphi^{-1}$. The Isomorphism Axiom requires that the combined map $\varphi \times \psi\colon A \times_\lozenge B \to B \times_\lozenge A$ preserve orientation, so we must have $(\varphi \times \psi)(A \times_\lozenge B) = B \times_\lozenge A$. This time, when we substitute in from Equation 9-5, we find that

$$B \times_\lozenge A = (\varphi \times \psi)\big(\eta\langle(\alpha, 0), (0, \beta)\rangle\big) = \eta\langle(\beta, 0), (0, \alpha)\rangle. \tag{9-7}$$

But Equations 9-6 and 9-7 are flatly contradictory. $\square$

In fact, this result is not really about fiber products, but rather about direct products, the direct product $A \times B$ being just that special case of the fiber product

$A \times B = A \times_\lozenge B$ in which the base space is $\lozenge$. Thus, Proposition 9-4 actually proves that there is no rule for orienting the direct products of oriented linear spaces that is both commutative and respects isomorphisms.

So we aren't going to get full commutativity. To enable our application to CAGD and robotics, commutativity had better hold for equidimensional problem instances. But that is already guaranteed without any need for additional axioms, as we saw in Exercise 9-3.

9.1.9 The Concatenate Axiom

Consider a binary problem instance of the form $A \xrightarrow{f} \lozenge \xleftarrow{g} B$. The factor maps f and g must both be identically zero, so the fiber product will be the entire direct product: $A \times_\lozenge B = \pm(A \times B)$. If our axioms are going to be complete, we need some axiom that determines the proper orientation for such a fiber product. We have already chosen to use the Concatenate Rule to orient direct products. The most straightforward thing to do is to use that same rule also for these fiber products. So, the *Concatenate Axiom* requires that $A \times_\lozenge B = A \times B$, where the direct product on the right-hand side is oriented using the Concatenate Rule. That is, the positive basis for A precedes the positive basis for B in assembling a positive basis for $A \times_\lozenge B$.

There is some unavoidable arbitrariness about this choice. If we liked, we could instead adopt the *Concatenate-Backwards Axiom*, which says that $A \times_\lozenge B = B \times A = (-1)^{ab}(A \times B)$. We simply have to choose something, and either of those choices turns out to be consistent with our other axioms.

Exercise 9-8 Suppose we take the base space S to be the direct product of the oriented spaces A and B, oriented via the Concatenate Rule — so we have $S = A \times B$. And suppose that we take our factor maps to be the standard injections; so $f\colon A \to A \times B$ is given by $f(\alpha) := (\alpha, 0)$, while $g\colon B \to A \times B$ is given by $g(\beta) := (0, \beta)$. Those two factor maps are always transverse, and the fiber product $A \times_S B = A \times_{(A\times B)} B$ is always zero-dimensional, representing the transverse intersection of the A and B axes in the direct product $A \times B$, as in Section 4.7.2. Is that intersection oriented positively or negatively?

Answer: It turns out that $A \times_{(A\times B)} B = (-1)^{ab}\lozenge$.

This result might lead some readers to suspect that we are making a mistake by adopting the Concatenate Axiom. The Concatenate-Backwards Axiom would eliminate the factor of $(-1)^{ab}$ in this formula; would that be the better choice? No. To see why not, try extending the example in this exercise from a binary intersection to a ternary one. With $S := A \times B \times C$ as our base space, we need to take the pairwise products $A \times B$, $B \times C$, and $A \times C$ as our factor spaces, rather than the individual spaces A, B, and C. But in which order should we put those three products? Surely $B \times C$ should go first, since it is the factor space in which A is missing. Indeed, after adopting the Concatenate Axiom, we shall find in Exercise 10-21 that the fiber product $(B \times C) \times_S (A \times C) \times_S (A \times B)$, which is zero-dimensional, is always positively oriented. The general, n-ary formula is

$$(\hat{A}_1 \times \cdots \times A_n) \times_S (A_1 \times \hat{A}_2 \times \cdots \times A_n) \times_S \cdots \times_S (A_1 \times \cdots \times \hat{A}_n) = +\lozenge,$$

where $S = A_1 \times \cdots \times A_n$ and the hats indicate omitted factors. Going back to the binary case $n = 2$, we should have considered the reversed product $B \times_S A = B \times_{(A\times B)} A$, since B is the factor space in which A is missing; and, under the Concatenate Axiom, that product is the simple one: $B \times_{(A\times B)} A = \lozenge$.

9.2 Completeness

The axioms in our list, while they are far from independent, are both consistent and complete. We first show their completeness.

Lemma 9-9 *Let $M \xrightarrow{p} U \xleftarrow{q} N$ be a binary problem instance; so the maps p and q are transverse and the spaces M, U, and N are oriented. Let L be any oriented linear space, and recall, from Section 5.4, that the map $0 \mathbin{\#} p\colon L \times M \to U$ is defined by $(0 \mathbin{\#} p)(\lambda, \mu) = 0(\lambda) + p(\mu) = p(\mu)$. If we assume the Isomorphism, Mixed Associativity and Concatenate Axioms, we then have*

$$L \times \bigl(M[p] \times_U [q]N\bigr) = (L \times M)[0 \mathbin{\#} p] \times_U [q]N,$$

as an equality between oriented linear spaces.

Proof We first apply the Axiom of Mixed Associativity to the ternary problem instance $L \xrightarrow{0} \lozenge \xleftarrow{0} M \xrightarrow{p} U \xleftarrow{q} N$, getting

$$L[0] \times_\lozenge [0]\bigl(M[p] \times_U [q]N\bigr) = \bigl(L[0] \times_\lozenge [0]M\bigr)[p'] \times_U [q]N,$$

where $p'\colon L \times_\lozenge M \to U$ is the map given by $p'(\lambda, \mu) = p(\mu)$. We then apply the Concatenate Axiom, on each side, to replace the fiber products over $\lozenge$ with direct products, noting that the map p' then becomes the map $0 \mathbin{\#} p$. (We need the Isomorphism Axiom, on the right-hand side, to ensure that the outer fiber product doesn't flip to the opposite orientation when we apply the Concatenate Axiom to its left factor space.) $\square$

Theorem 9-10 *There is at most one orientation rule for transverse, binary fiber products of oriented linear spaces that satisfies these five axioms: Isomorphism, Stability, Left Identity, Mixed Associativity, and Concatenate.*

Proof Given any binary problem instance $A \xrightarrow{f} S \xleftarrow{g} B$, we claim that the orientation of the fiber product $A[f] \times_S [g]B$ is determined by the listed axioms.

We first apply Lemma 9-9 to the original instance $A \xrightarrow{f} S \xleftarrow{g} B$ with $L := S$, learning that

$$S \times \big(A[f] \times_S [g]B\big) = (S \times A)[0 \mathbin{\#} f] \times_S [g]B.$$

The fiber product on the left in this equality is the arbitrary one that we started with, while the fiber product on the right has a left factor, $S \times A$, whose dimension is at least the dimension of the base space S. Hence, it will suffice for us to demonstrate that all binary problem instances $A \xrightarrow{f} S \xleftarrow{g} B$ in which $a \geq s$ are determined. (In fact, we could restrict ourselves further to binary problem instances in which $a \geq s$ and the values $f(\alpha)$ of the left-hand factor map don't depend upon the first s coordinates of the vector α in A. But we won't need that additional restriction.)

In any problem instance with $a \geq s$, there always exist linear surjections from A onto S, so let $h\colon A \to S$ be one such. There are only finitely many real numbers t for which the linear combination $f + th$ will fail to be a surjection. Hence, perturbing the first factor map to be $f_t := f + th$ while leaving the second factor map alone, $g_t := g$, will guarantee transversality for all real numbers t that are sufficiently close to zero: $\mathrm{Im}(f_t) + \mathrm{Im}(g_t) = S + \mathrm{Im}(g) = S$. Thus, our current problem instance $A \xrightarrow{f} S \xleftarrow{g} B$ is the limit of problem instances $A \xrightarrow{f_t} S \xleftarrow{g_t} B$ in which the first factor map f_t is surjective. By the Stability Axiom, it suffices to show that all instances of this latter type are determined.

So suppose that the map f in the problem instance $A \xrightarrow{f} S \xleftarrow{g} B$ is surjective. Let C be some complement of $\mathrm{Ker}(f)$ in A, viewed for now as unoriented; we'll orient C in a moment. The linear map f gives us a bijection from $A/\mathrm{Ker}(f)$ to $\mathrm{Im}(f) = S$ and hence also a bijection from C to S; call that latter bijection $\bar{f}\colon C \to S$, and orient C so that the bijection $\bar{f}$ preserves orientation. And let K denote the linear space $\mathrm{Ker}(f)$, oriented so that $K \times C = A$ is an equality between oriented linear spaces. We now apply Lemma 9-9 to the binary problem instance

$$C \xrightarrow{\bar{f}} S \xleftarrow{g} B,$$

with $L := K$, learning that

$$K \times \big(C[\bar{f}] \times_S [g]B\big) = (K \times C)[0 \mathbin{\#} \bar{f}] \times_S [g]B.$$

The right-hand side is the fiber product $A[f] \times_S [g]B$ whose orientation we are trying to show determined. On the left-hand side, the left-hand factor map of the fiber product is the orientation-preserving bijection $\bar{f}$. Hence, by the Isomorphism and Left Identity Axioms, we have

$$K \times \big(C[\bar{f}] \times_S [g]B\big) = K \times (S[1] \times_S [g]B) = K \times B.$$

Thus, the orientations of all transverse, binary fiber products are determined by the listed axioms. $\square$

9.3 Consistency: Calibrating the Delta Rule

To show that our axioms are consistent, we need to construct an orientation rule that satisfies them all. We shall construct that rule by adding an explicit calibration factor to the Delta Rule of Definition 7-2. By the way, the following version of the Delta Rule has also been generalized, in the obvious way, from pure to mixed fiber products.

Definition 9-11 (The Calibrated Delta Rule) For some $n \geq 1$, let

$$A_1 \xrightarrow{f_1} S_1 \xleftarrow{g_2} A_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} S_{n-1} \xleftarrow{g_n} A_n$$

be an n-ary zigzag, consisting of oriented linear spaces and linear maps. Let P be the mixed fiber product $P := A_1[f_1] \times_{S_1} [g_2]A_2[f_2] \times_{S_2} \cdots \times_{S_{n-1}} [g_n]A_n$, which we view as a subspace of the direct product $D := A_1 \times \cdots \times A_n$. We define the difference map $\Delta\colon D \to S_1 \times \cdots \times S_{n-1}$ by

$$\Delta(\alpha_1, \ldots, \alpha_n) := \big(g_2(\alpha_2) - f_1(\alpha_1),\ \ldots,\ g_n(\alpha_n) - f_{n-1}(\alpha_{n-1})\big).$$

We assume that our zigzag is transverse, which means that the difference map Δ is surjective. If C denotes any oriented complement of $P = \mathrm{Ker}(\Delta)$ in D, it follows that Δ maps C bijectively onto $\mathrm{Im}(\Delta) = S_1 \times \cdots \times S_{n-1}$. The Calibrated Delta Rule orients the fiber product P so that

$$\text{(9-12)} \qquad \frac{P \oplus C}{D} = \frac{(A_1 \times_{S_1} \cdots \times_{S_{n-1}} A_n) \oplus C}{A_1 \times \cdots \times A_n} = (-1)^{\kappa} \frac{\Delta(C)}{S_1 \times \cdots \times S_{n-1}}.$$

The calibration is determined by the exponent κ in the factor $(-1)^{\kappa}$. We choose our calibration factor to depend upon the dimensions $(a_1, \ldots, a_n)$ of the factor spaces and the dimensions $(s_1, \ldots, s_{n-1})$ of the base spaces as follows: $\kappa := \kappa_n(a_1, \ldots, a_n; s_1, \ldots, s_{n-1})$, where the function $\kappa_n\colon \mathbb{N}^n \times \mathbb{N}^{n-1} \to \mathbb{Z}/2\mathbb{Z}$ is defined by

$$\text{(9-13)} \qquad \kappa_n(a_1, \ldots, a_n; s_1, \ldots, s_{n-1}) := \sum_{1 \leq i \leq j < n} s_i(a_{j+1} - s_j).$$

The main new idea in Definition 9-11 is the calibration function κ_n given in Equation 9-13, which is rather mysterious at first glance. Of course, we made various arbitrary choices in setting up the Delta Rule; if we changed our minds about those choices, then the proper calibration function would change as well. But the formula for κ_n is so wild that no simple readjustment of our arbitrary choices could possibly free us entirely from the need to calibrate.

We can avoid the issue of calibration in the equidimensional case, however; that is another way in which the equidimensional case is easy. For any integer m, we have

$$\kappa_n(\underbrace{m, \ldots, m}_{n}; \underbrace{m, \ldots, m}_{n-1}) = 0,$$

since each term $s_i(a_{j+1} - s_j) = m(m - m)$ in the sum is zero. Indeed, when the factor maps g_2 through g_n are all invertible, that by itself is enough to guarantee that $a_{j+1} = s_j$ for j from 1 to $n - 1$, and hence $\kappa_n(a_1, \ldots, a_n; s_1, \ldots, s_{n-1}) = 0$. This explains why we found, in proving Theorem 7-12, that the Uncalibrated Delta Rule already obeys the case $k = 1$ of the All-but-One Invertible Law 4-7.

Our next goal is to verify that the axioms in Section 9.1 are consistent by showing that the Delta Rule, when calibrated by the particular function κ_n given in Equation 9-13, satisfies all of those axioms. Of course, those axioms talk only about the binary case, so we shall typically work only with the binary calibration function $\kappa_2 \colon \mathbb{N}^2 \times \mathbb{N} \to \mathbb{Z}/2\mathbb{Z}$ given by $\kappa_2(a, b; s) = s(b - s)$. In the process of verifying associativity, however, we shall go further and verify that κ_n, as given in Equation 9-13, is the proper calibration function for all arities n.

Theorem 9-14 *Let the* Binary Delta Rule *be the binary case of the Delta Rule, calibrated by letting $\kappa_2(a, b; s) := s(b - s)$. The Binary Delta Rule orients every transverse fiber product of linear spaces in a way that satisfies all of our axioms: Isomorphism; Stability; Left, Right, and Base Reversing; Both, Left, and Right Identity; Mixed Associativity; and Concatenate.*

Because of associativity, the Binary Delta Rule extends uniquely to an n-ary Delta Rule, for all n; and that rule corresponds to adopting the calibration function

$$\kappa_n(a_1, \ldots, a_n; s_1, \ldots, s_{n-1}) := \sum_{1 \le i \le j < n} s_i(a_{j+1} - s_j),$$

given in Equation 9-13.

We tackle the proof of Theorem 9-14 one axiom at a time.

9.3.1 The Isomorphism Axiom

The Isomorphism Axiom holds regardless of how the Delta Rule is calibrated, and the proof is quite straightforward. We start with two binary problem instances $A \xrightarrow{f} S \xleftarrow{g} B$ and $A' \xrightarrow{f'} S' \xleftarrow{g'} B'$ that are isomorphic. If C is any oriented complement of $P := A \times_S B$ in $A \times B$, then the image of C under the isomorphism, call it C', will be an oriented complement of $P' := A' \times_{S'} B'$ in $A' \times B'$. With this choice for C', the right-hand quotients in the equations

$$\frac{P \oplus C}{A \times B} = (-1)^{\kappa_2(a,b;s)} \frac{\Delta(C)}{S} \qquad \text{and} \qquad \frac{P' \oplus C'}{A' \times B'} = (-1)^{\kappa_2(a',b';s')} \frac{\Delta'(C')}{S'}$$

will have the same sign. The calibration factors are also the same, because we have $a = a'$, $b = b'$, and $s = s'$. So, on the left-hand side, the isomorphism must carry P to $+P'$ rather than to $-P'$, as the Isomorphism Axiom requires.

9.3.2 The Stability Axiom

We saw, in Proposition 7-6, that the Uncalibrated Delta Rule is stable and that it remains stable under any recalibration that depends only upon the dimensions of the factor spaces and base spaces.

9.3.3 The Reversing Axioms

The three Reversing Axioms also hold, no matter what calibration function we adopt. Consider the relevant formula:

$$\frac{P \oplus C}{A \times B} = (-1)^{\kappa_2(a,b;s)} \frac{\Delta(C)}{S}.$$

Reversing the orientation on the first factor space A reverses the denominator on the left-hand side and nothing else; so we have to reverse P also, to keep the equation valid. Reversing B is similar. Reversing S reverses the denominator on the right-hand side and nothing else; so, again, we have to reverse P. Thus, any calibration of the Delta Rule satisfies all three Reversing Axioms.

9.3.4 The Both Identities Axiom

The Both Identities Axiom is the first of our axioms that actually constrains the calibration function. It turns out that the Both Identities Axiom holds just when κ_2 is zero on its main diagonal — that is, when $\kappa_2(m, m; m) = 0$, for all $m \geq 0$. Our chosen calibration $\kappa_2(a, b; s) = s(b - s)$ certainly meets this requirement.

Here is why. For the problem instance $S \xrightarrow{1} S \xleftarrow{1} S$, the Delta Rule says:

$$\frac{P \oplus C}{S \times S} = (-1)^{\kappa_2(s,s;s)} \frac{\Delta(C)}{S} \tag{9-15}$$

The fiber product $P = S \times_S S$ is, up to orientation, the main diagonal of $S \times S$, the space $P = \{(\sigma, \sigma) \mid \sigma \in S\}$. So we have $P = \eta\,(1_{S\leftarrow S}, 1_{S\leftarrow S})(S)$ for some sign $\eta = \pm 1$, and the Both Identities Axiom requires that $\eta = +1$.

We get to choose C to be any oriented complement of P in $S \times S$, and it is convenient to choose $C = (0, S)$, that is, to set $C := (0_{S\leftarrow S}, 1_{S\leftarrow S})(S)$. With this choice, since the difference map Δ has the matrix $\Delta = \begin{pmatrix} -1_{S\leftarrow S} & 1_{S\leftarrow S} \end{pmatrix}$, we have

$$\Delta(C) = \begin{pmatrix} -1_{S\leftarrow S} & 1_{S\leftarrow S} \end{pmatrix} \begin{pmatrix} 0_{S\leftarrow S} \\ 1_{S\leftarrow S} \end{pmatrix} (S) = +S.$$

Thus, the quotient on the right-hand side of Equation 9-15 reduces to $+\Diamond$.

On the left-hand side, pasting together our recipes for P and for C as specified in Equation 5-7, we find that

$$P \oplus C = \eta \begin{pmatrix} 1_{S\leftarrow S} & 0_{S\leftarrow S} \\ 1_{S\leftarrow S} & 1_{S\leftarrow S} \end{pmatrix} (S \times S).$$

Since either an elementary row or column operation converts this matrix into the identity, we conclude that $(P \oplus C)/(S \times S) = \eta\diamondsuit$. Thus, to make Equation 9-15 hold with $\eta = +1$, it suffices to ensure that $\kappa_2(s, s; s) = 0$.

9.3.5 The Left Identity Axiom

To analyze the Left Identity Axiom, we consider the instance $S \xrightarrow{1} S \xleftarrow{g} B$, for which the Delta Rule says:

$$\frac{P \oplus C}{S \times B} = (-1)^{\kappa_2(s,b;s)} \frac{\Delta(C)}{S} \tag{9-16}$$

The fiber product $P = S \times_S B$ here is the subspace of $A \times B = S \times B$ consisting of all pairs of the form $(g(\beta), \beta)$, for β in B. We thus have $P = \eta(g, 1_{B\leftarrow B})(B)$ for some sign $\eta = \pm 1$, and the Left Identity Axiom requires $\eta = +1$.

We choose C, this time, to be the space $C := (S, 0) = (1_{S\leftarrow S}, 0_{B\leftarrow S})(S)$. With that choice, what happens on the right-hand side? The difference map Δ here has the matrix $\Delta = \begin{pmatrix} -1_{S\leftarrow S} & g \end{pmatrix}$, so we have

$$\Delta(C) = \begin{pmatrix} -1_{S\leftarrow S} & g \end{pmatrix} \begin{pmatrix} 1_{S\leftarrow S} \\ 0_{S\leftarrow S} \end{pmatrix} (S) = \neg(S) = (-1)^s S,$$

where "$\neg$" is the negation map discussed in Exercise 5-3. We conclude that $\Delta(C)/S = \neg(S)/S = (-1)^s\diamondsuit$. Since k^2 is odd just when k is odd, we can rewrite this as $(-1)^{s^2}\diamondsuit$.

As for the left-hand quotient in Equation 9-16, pasting together the column matrices that define P and C, we have

$$P \oplus C = \eta \begin{pmatrix} g & 1_{S\leftarrow S} \\ 1_{B\leftarrow B} & 0_{B\leftarrow S} \end{pmatrix} (B \times S).$$

To simplify this matrix, we multiply the second column by g on the right and subtract from the first column — or we multiply the second row by g on the left and subtract from the first row. Either way, we conclude that

$$P \oplus C = \eta \begin{pmatrix} 0_{S\leftarrow B} & 1_{S\leftarrow S} \\ 1_{B\leftarrow B} & 0_{B\leftarrow S} \end{pmatrix} (B \times S),$$

where this matrix denotes simply the swapping map from $B \times S$ to $S \times B$. From our study of the Concatenate Rule in Section 5.2, we conclude that $P \oplus C = \eta(-1)^{bs}(S \times B)$, so the left-hand quotient is $\eta(-1)^{bs}\diamondsuit$.

To make Equation 9-16 hold with $\eta = +1$, as the Left Identity Axiom requires, we must have

$$(-1)^{bs} = (-1)^{\kappa_2(s,b;s)}(-1)^{s^2}.$$

So the calibration function κ_2 must satisfy $\kappa_2(s, b; s) = bs - s^2 = s(b - s)$. And our proposed calibration, given by $\kappa_2(a, b; s) = s(b - s)$, meets this requirement.

9.3.6 The Right Identity Axiom

The Right Identity Axiom is simpler. For the instance $A \xrightarrow{f} S \xleftarrow{1} S$, the Delta Rule says

$$\frac{P \oplus C}{A \times S} = (-1)^{\kappa_2(a,s;s)} \frac{\Delta(C)}{S},$$

where the fiber product is the subspace $P = +(1_{A \leftarrow A}, f)(A)$ of $A \times S$ and the Right Identity Axiom requires the positive sign. Choosing $C := (0, S) = (0_{A \leftarrow S}, 1_{S \leftarrow S})(S)$ reduces the quotient on the right-hand side to $+\Diamond$, while the direct sum $P \oplus C$ is the image of $A \times S$ under the map with matrix

$$\begin{pmatrix} 1_{A \leftarrow A} & 0_{A \leftarrow S} \\ f & 1_{S \leftarrow S} \end{pmatrix}.$$

Using an elementary row or column operation to zero out the f entry, the left-hand quotient also reduces to $+\Diamond$. So the Right Identity Axiom holds just when the calibration function κ_2 satisfies the identity $\kappa_2(a, s; s) = 0$. And our proposed function $\kappa_2(a, b; s) := s(b - s)$ meets this requirement also.

9.3.7 The Axiom of Mixed Associativity

Verifying the Axiom of Mixed Associativity involves rather intricate reasoning, but the challenges are more notational than conceptual.

The axiom itself, $(A \times_S B) \times_T C = A \times_S (B \times_T C)$, involves only binary fiber products. One way to show that the two sides have the same orientation is to show that each side has the same orientation as the ternary product $A \times_S B \times_T C$. We can unify the arguments for the left and right sides by generalizing as follows. Consider the nested product

$$A_1 \times_{S_1} \cdots \times_{S_{i-2}} A_{i-1} \times_{S_{i-1}} (A_i \times_{S_i} A_{i+1}) \times_{S_{i+1}} A_{i+2} \times_{S_{i+2}} \cdots \times_{S_{n-1}} A_n,$$

where $1 \le i < n$. That is, we first combine the adjacent factors A_i and A_{i+1} in an inner, binary fiber product. We then combine the result of that inner product with the remaining factors, using an outer fiber product of arity $n - 1$. We shall show that the orientation that results from this two-step process agrees with the orientation that would result from computing the overall product in one n-ary step. Once we have that general result, the two special cases $(i, n) := (1, 3)$ and $(i, n) := (2, 3)$ combine to establish the Axiom of Mixed Associativity.

In order for the nested orientation and the overall orientation to be defined, we must calibrate the Delta Rule for products of arities $n - 1$ and n. As we claimed in Section 9.3, the proper calibration function for any positive n is

$$\kappa_n(a_1, \ldots, a_n; s_1, \ldots, s_{n-1}) := \sum_{1 \le k \le l < n} s_k(a_{l+1} - s_l).$$

In addition to establishing associativity, our argument below verifies that these are the proper calibration functions for products of arity $n > 2$. In what follows, it is helpful to write κ_n with the sum on k outside:

$$\kappa_n(a_1, \ldots, a_n; s_1, \ldots, s_{n-1}) := \sum_{1\le k<n} s_k \sum_{k\le l<n} (a_{l+1} - s_l)$$

To make our formulas somewhat shorter, let's use the symbol I to denote the spaces associated with the inner fiber product, while H and J denote the spaces to its left and to its right. That is, for the factor spaces (A_k), we set:

$$\begin{aligned} A_H &:= A_1 \times \cdots \times A_{i-1} \\ A_I &:= A_i \times A_{i+1} \\ A_J &:= A_{i+2} \times \cdots \times A_n \end{aligned}$$

We also abbreviate the full direct product $A_1 \times \cdots \times A_n = A_H \times A_I \times A_J$ as A_{HIJ}. For the base spaces (S_k), we do something similar, except that one less space is involved:

$$\begin{aligned} S_H &:= S_1 \times \cdots \times S_{i-1} \\ S_I &:= S_i \\ S_J &:= S_{i+1} \times \cdots \times S_{n-1} \\ S_{HIJ} &:= S_1 \times \cdots \times S_{n-1} \end{aligned}$$

We first tackle the inner fiber product $A_i \times_{S_i} A_{i+1}$, which we denote by P_I. Let C_I be some oriented complement of P_I in the direct product $A_I = A_i \times A_{i+1}$. Applying the Delta Rule to the inner product, we have

$$\text{(Inner)} \qquad \frac{P_I \oplus C_I}{A_I} = (-1)^{s_i(a_{i+1}-s_i)} \frac{\Delta_I(C_I)}{S_I},$$

where $\Delta_I \colon A_I \to S_I$ is the difference map that is given by $\Delta_I(\alpha_i, \alpha_{i+1}) := g_{i+1}(\alpha_{i+1}) - f_i(\alpha_i)$. Call that the Inner Equation.

Now for the outer fiber product $A_1 \times_{S_1} \cdots \times_{S_{i-2}} A_{i-1} \times_{S_{i-1}} P_I \times_{S_{i+1}} A_{i+2} \times_{S_{i+2}} \cdots \times_{S_{n-1}} A_n$, which we denote by P_O. Let C_O be some oriented complement of P_O in the corresponding direct product $A_1 \times \cdots \times A_{i-1} \times P_I \times A_{i+2} \times \cdots \times A_n = A_H \times P_I \times A_J$. With our abbreviations, the Delta Rule applied to the outer product says that

$$\text{(Outer)} \qquad \frac{P_O \oplus C_O}{A_H \times P_I \times A_J} = (-1)^{K_O} \frac{\Delta_O(C_O)}{S_H \times S_J},$$

where $\Delta_O \colon A_H \times P_I \times A_J \to S_H \times S_J$ is the map defined by

$$\begin{aligned} &\Delta_O(\alpha_1, \ldots, \alpha_{i-1}, (\alpha_i, \alpha_{i+1}), \alpha_{i+2}, \ldots, \alpha_n) := \\ &\qquad \big(g_2(\alpha_2) - f_1(\alpha_1), \ldots, g_i(\alpha_i) - f_{i-1}(\alpha_{i-1}), \\ &\qquad\quad g_{i+2}(\alpha_{i+2}) - f_{i+1}(\alpha_{i+1}), \ldots, g_n(\alpha_n) - f_{n-1}(\alpha_{n-1})\big). \end{aligned}$$

Note that the difference $g_{i+1}(\alpha_{i+1}) - f_i(\alpha_i)$, which is identically zero on P_I, is omitted when defining the outer difference map Δ_O. In this Outer Equation, the exponent K_O of the calibration factor can be written

$$K_O := \sum_{1\le k<i} s_k \Bigg(\sum_{k\le l<i-1} (a_{l+1} - s_l) + (p_I - s_{i-1}) + \sum_{i<l<n} (a_{l+1} - s_l) \Bigg) + \sum_{i<k<n} s_k \sum_{k\le l<n} (a_{l+1} - s_l),$$

where $p_I := \dim(P_I)$. Since the inner fiber product is transverse, we have $p_I = a_i + a_{i+1} - s_i$; so this simplifies to

$$K_O = \sum_{\substack{1\le k<n \\ k\neq i}} s_k \sum_{k\le l<n} (a_{l+1} - s_l).$$

Finally, let $P_V := A_1 \times_{S_1} \cdots \times_{S_{n-1}} A_n$ be the overall fiber product, viewed as a subspace of the overall direct product $A_1 \times \cdots \times A_n = A_{HIJ}$. We need some oriented complement C_V of P_V in A_{HIJ}. We choose to set $C_V := C_O \oplus C_I'$, where $C_I' := 0 \times C_I \times 0$ is C_I, viewed as subset of A_{HIJ}. Applying the Delta Rule to the overall product then gives us

$$\text{(Overall)} \qquad \frac{P_V \oplus C_O \oplus C_I'}{A_{HIJ}} = (-1)^{K_V} \frac{\Delta_V(C_O \oplus C_I')}{S_{HIJ}}.$$

In this Overall Equation, the difference map Δ_V is the obvious one, while the exponent K_V of the calibration factor is

$$K_V := \sum_{1\le k<n} s_k \sum_{k\le l<n} (a_{l+1} - s_l).$$

The spaces P_V and P_O are the same, except possibly for their orientations; and we want to show that their orientations are the same as well. Let the sign $\eta = \pm 1$ be such that $P_V = \eta P_O$; we shall show that the Inner, Outer, and Overall Equations imply that $\eta = +1$.

We begin with the quotient on the left-hand side of the Overall Equation, which we'll call the Overall left quotient:

$$\text{(9-17)} \qquad \frac{P_V \oplus C_O \oplus C_I'}{A_{HIJ}} = \eta \frac{P_O \oplus C_O \oplus C_I'}{A_{HIJ}}$$

We can produce a second quotient with the same numerator by starting with the Outer left quotient and adding C_I' as a direct summand at the right end of both the numerator and denominator:

$$\begin{aligned} \frac{P_O \oplus C_O}{A_H \times P_I \times A_J} &= \frac{P_O \oplus C_O \oplus C_I'}{(A_H \times P_I \times A_J) \oplus C_I'} \\ &= \frac{P_O \oplus C_O \oplus C_I'}{(A_H \times P_I \times A_J) \oplus (0 \times C_I \times 0)} \\ &= (-1)^{\dim(C_I)\dim(A_J)} \frac{P_O \oplus C_O \oplus C_I'}{A_H \times (P_I \oplus C_I) \times A_J} \end{aligned}$$

Since the inner fiber product is transverse, we have $\dim(C_I) = s_i$; and we trivially have $\dim(A_J) = a_{i+2} + \cdots + a_n$. We now rewrite Equation 9-17 with the right-hand quotient expanded as the product of two quotients, and we use what we've just learned to simplify the first of them:

$$\begin{aligned}\frac{P_V \oplus C_O \oplus C_I'}{A_{HIJ}} &= \eta \frac{P_O \oplus C_O \oplus C_I'}{A_H \times (P_I \oplus C_I) \times A_J} \times \frac{A_H \times (P_I \oplus C_I) \times A_J}{A_{HIJ}} \\ &= \eta(-1)^{s_i(a_{i+2}+\cdots+a_n)} \frac{P_O \oplus C_O}{A_H \times P_I \times A_J} \times \frac{A_H \times (P_I \oplus C_I) \times A_J}{A_{HIJ}}\end{aligned}$$

The rightmost quotient in this expansion also arises if we start with the Inner left quotient and add direct-product factors of A_H on the left and A_J on the right of both the numerator and denominator:

$$\frac{P_I \oplus C_I}{A_I} = \frac{A_H \times (P_I \oplus C_I) \times A_J}{A_H \times A_I \times A_J} = \frac{A_H \times (P_I \oplus C_I) \times A_J}{A_{HIJ}}$$

So we deduce the following relationship between the Overall, Outer, and Inner left quotients:

$$\frac{P_V \oplus C_O \oplus C_I'}{A_{HIJ}} = \eta(-1)^{s_i(a_{i+2}+\cdots+a_n)} \frac{P_O \oplus C_O}{A_H \times P_I \times A_J} \times \frac{P_I \oplus C_I}{A_I}$$

Substituting the right-hand sides of the Overall, Outer, and Inner Equations for their left-hand sides, we find that

(9-18)

$$\begin{aligned}(-1)^{K_V} \frac{\Delta_V(C_O \oplus C_I')}{S_{HIJ}} &= \eta(-1)^{s_i(a_{i+2}+\cdots+a_n)+K_O} \frac{\Delta_O(C_O)}{S_H \times S_J} (-1)^{s_i(a_{i+1}-s_i)} \frac{\Delta_I(C_I)}{S_I} \\ &= \eta(-1)^{s_i(a_{i+2}+\cdots+a_n)+K_O+s_i(a_{i+1}-s_i)} \frac{\Delta_O(C_O) \times \Delta_I(C_I)}{(S_H \times S_J) \times S_I}\end{aligned}$$

Note that any point $(\alpha_1, \ldots, \alpha_{i-1}, (\alpha_i, \alpha_{i+1}), \alpha_{i+2}, \ldots, \alpha_n)$ in C_O has $f_i(\alpha_i) = g_{i+1}(\alpha_{i+1})$, since C_O is a subset of $A_H \times P_I \times A_J$. Thus, the S_I coordinate in $\Delta_V(C_O)$, the coordinate that is omitted in $\Delta_O(C_O)$, is always 0. So we can rewrite the quotient on the right-hand side of Equation 9-18 using direct sums as

$$\frac{\Delta_O(C_O) \times \Delta_I(C_I)}{(S_H \times S_J) \times S_I} = \frac{\Delta_V(C_O) \oplus (0 \times \Delta_I(C_I) \times 0)}{(S_H \times 0 \times S_J) \oplus (0 \times S_I \times 0)}.$$

Compare this with the quotient on the left-hand side of Equation 9-18, which is

$$\frac{\Delta_V(C_O \oplus C_I')}{S_{HIJ}} = \frac{\Delta_V(C_O) \oplus \Delta_V(C_I')}{S_{HIJ}}.$$

The second summand in each of these numerators is an oriented complement of the common first summand $\Delta_V(C_O)$ in S_{HIJ}, so Exercise 5-5 is relevant here. Let

$\phi_i\colon S_{HIJ} \to S_I$ denote projection onto the factor space $S_I = S_i$, and note that $\mathrm{Ker}(\phi_i) = S_H \times 0 \times S_J = \pm\Delta_V(C_O)$. Exercise 5-5 then tells us that our two numerators $\Delta_V(C_O) \oplus (0 \times \Delta_I(C_I) \times 0)$ and $\Delta_V(C_O) \oplus \Delta_V(C_I')$ will give the same orientation to S_{HIJ} just when the two expressions $\phi_i(0 \times \Delta_I(C_I) \times 0)$ and $\phi_i(\Delta_V(C_I'))$ give the same orientation to S_i. And they do: Consider any point (α_i, α_{i+1}) in C_I. This point maps to

$$(\underbrace{0,\ldots,0}_{i-1}, g_{i+1}(\alpha_{i+1}) - f_i(\alpha_i), \underbrace{0,\ldots,0}_{n-i-1}) \qquad \text{in } (0 \times \Delta_I(C_I) \times 0) \text{ and to}$$

$$(\underbrace{0,\ldots,0}_{i-2}, g_i(\alpha_i), g_{i+1}(\alpha_{i+1}) - f_i(\alpha_i), -f_{i+1}(\alpha_{i+1}), \underbrace{0,\ldots,0}_{n-i-2}) \qquad \text{in } \Delta_V(C_I'),$$

and both of those map to $g_{i+1}(\alpha_{i+1}) - f_i(\alpha_i)$ under ϕ_i. So we conclude that

$$\frac{\Delta_V(C_O \oplus C_I')}{S_{HIJ}} = \frac{\Delta_O(C_O) \times \Delta_I(C_I)}{S_{HIJ}}.$$

Comparing this with Equation 9-18, we next account for the difference between the denominators S_{HIJ} and $(S_H \times S_J) \times S_I = (S_H \times 0 \times S_J) \oplus (0 \times S_I \times 0)$. We have

$$S_{HIJ} = S_H \times S_I \times S_J = (-1)^{\dim(S_I)\dim(S_J)}\big((S_H \times 0 \times S_J) \oplus (0 \times S_I \times 0)\big),$$

where $\dim(S_I)\dim(S_J) = s_i(s_{i+1} + \cdots + s_{n-1})$. Equation 9-18 thus reduces to the following equation about signs:

$$(-1)^{K_V} = \eta(-1)^{s_i(a_{i+2}+\cdots+a_n)+K_O+s_i(a_{i+1}-s_i)+s_i(s_{i+1}+\cdots+s_{n-1})}.$$

The exponent K_O cancels out most of K_V, leaving just the terms in K_V with $k = i$, which are $s_i \sum_{i \le l < n}(a_{l+1} - s_l)$. Recalling that signs don't matter in dealing with exponents of -1, those terms exactly cancel against the remaining explicit terms in the exponent on the right-hand side. We conclude, from all of this, that $\eta = +1$, which completes our proof of mixed associativity.

9.3.8 The issue of commutativity

While we know that we can't achieve full commutativity, it is interesting to see what does happen about commutativity, with our chosen calibration function κ_n. We claim that $A \times_S B = (-1)^{(a-s)(b-s)}(B \times_S A)$. Recall that direct products skew commute just when the dimensions of both factor spaces are odd. Fiber products generalize this by skew commuting just when both factor spaces have dimensional parity opposite to that of the base space.

This claim is easy to verify. Let $P := A \times_S B$ and $P' := B \times_S A$ be the two fiber products, each oriented as specified by the Delta Rule. Let C be some oriented complement of P in $A \times B$, and let C' be the image of C under the swapping map $(\alpha, \beta) \mapsto (\beta, \alpha)$. The space C' is then an oriented complement of

P' in $B \times A$. If we apply that swapping map to the fiber product P, we get either P' or $-P'$, and our challenge is to determine which. Applying the Delta Rule to the unprimed case, we have

$$\frac{P \oplus C}{(A, 0) \oplus (0, B)} = (-1)^{s(b-s)} \frac{\Delta(C)}{S}.$$

In the primed case, we get

$$\frac{P' \oplus C'}{(B, 0) \oplus (0, A)} = (-1)^{s(a-s)} \frac{\Delta'(C')}{S}.$$

Noting that $\Delta'(\beta, \alpha) = f(\alpha) - g(\beta)$, we see that $\Delta'(C') = \neg(\Delta(C))$, where "$\neg$" is the negation map discussed in Exercise 5-3. As for the primed left quotient, let's apply the inverse swapping map $(\beta, \alpha) \mapsto (\alpha, \beta)$ to the direct summands in both numerator and denominator; this just changes our coordinate system and hence does not affect the sign of the quotient. So the primed case reduces to

$$\frac{(\pm P) \oplus C}{(0, B) \oplus (A, 0)} = (-1)^{s(a-s)} \frac{\neg(\Delta(C))}{S}.$$

Comparing this to the unprimed case, the left-hand denominators differ by $(-1)^{ab}$; the calibration factors differ by $(-1)^{sb-sa} = (-1)^{-sa-sb}$; and the right-hand numerators differ by $(-1)^s = (-1)^{s^2}$. So the proper sign in $(\pm P)$ is $(-1)^{(a-s)(b-s)}$, as we claimed.

9.3.9 The Concatenate Axiom

The Concatenate Axiom is easy. Suppose that $S = \lozenge$ is zero-dimensional and positively oriented; the Concatenate Axiom requires that $A \times_{\lozenge} B = A \times B$. To verify this, we apply the Delta Rule, getting

$$\frac{P \oplus C}{A \times B} = (-1)^{0(b-0)} \frac{\Delta(C)}{\lozenge}.$$

Noting that C is zero-dimensional, we can take it to be positively oriented. We then have $P \oplus C = P$ and $\Delta(C) = \Delta(\lozenge) = \lozenge$, so we have

$$\frac{P}{A \times B} = \frac{A \times_{\lozenge} B}{A \times B} = +\lozenge.$$

That being the last of the axioms on our list, the proof of Theorem 9-14 is now complete. So there is a unique rule for orienting transverse, mixed fiber products that satisfies all of the axioms on our list; that rule is the Delta Rule, calibrated as in Equation 9-13. In what follows, we shall refer to the orientations produced by that rule as the *Proper Orientations*. Furthermore, given any transverse, n-ary zigzag $A_1 \xrightarrow{f_1} S_1 \xleftarrow{g_2} A_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} S_{n-1} \xleftarrow{g_n} A_n$, we henceforth use the formula $A_1 \times_{S_1} A_2 \times_{S_2} \cdots \times_{S_{n-1}} A_n$ to denote its mixed fiber product, oriented properly.

(We defined the Proper Orientation for an all-invertible, pure fiber product back in Definition 7-10, and we extended that by continuity to all transverse, equidimensional, pure fiber products in Definition 7-15. We are now extending further, both to the any-dimensional case and from pure to mixed fiber products; but this new extension is consistent with those earlier definitions.)

Chapter 10

Adjusting the axioms

We have achieved our key goal: a consistent, complete, and compelling family of axioms for a rule that orients any-dimensional transverse fiber products. In particular, I am personally pleased that there turns out to be an orientation rule that satisfies the Axiom of Mixed Associativity.

In this final chapter, we consider several ways in which we might adjust our family of axioms. First, we study how free we would be to vary our orientation rule, were we to give up on certain of our axioms. Second, we discuss several ways in which our list of axioms could be shortened by combining a group of axioms into a single, more powerful identity.

10.1 Twisted orientation rules

The Proper Orientation Rule is the unique rule that satisfies all of our axioms. But suppose that we were willing to abandon some of our axioms. What freedom would we thereby acquire to adopt some orientation rule different from the Proper Rule? To study this issue, we introduce the concept of a "dimensional twist".

Given a binary problem instance $A \xrightarrow{f} S \xleftarrow{g} B$, we have agreed that the formula $A \times_S B$ denotes the fiber product with its Proper Orientation. Let's denote by $A \times_S^* B = \pm(A \times_S B)$ that same fiber product, but oriented by some possibly improper orientation rule.

For the problem instance $A \xrightarrow{f} S \xleftarrow{g} B$ to be transverse, we must have $a + b \geq s$; let $\mathcal{T}$ be the set of all triples of nonnegative integers $(a, b; s)$ with $a + b \geq s$. Define a *dimensional twist* to be any function $w\colon \mathcal{T} \to \mathbb{Z}/2\mathbb{Z}$; so, for each triple $(a, b; s)$ in $\mathcal{T}$, we get to choose a free bit $w(a, b; s)$. Given a dimensional twist w, the corresponding *twisted orientations* differ from the Proper Orientations as follows:

$$A \times_S^w B := (-1)^{w(a,b;s)}(A \times_S B).$$

What do our various axioms say about which dimensional twists are legal?

10.1.1 Twists are the only reasonable adjustments

It's easy to see that the Isomorphism, Stability, and Reversing Axioms hold for any twisted orientation rule, just as they do for the untwisted Proper Orientations. The more interesting result goes in the other direction.

Proposition 10-1 *Any orientation rule $\times^*$ that satisfies the Isomorphism, Stability, and Reversing Axioms is the twisted orientation rule $\times^* := \times^w$ corresponding to some dimensional twist $w\colon \mathcal{T} \to \mathbb{Z}/2\mathbb{Z}$.*

Proof Let $A \times^*_S B$ denote the result of orienting the fiber product $A \times_S B$ by some rule that satisfies the Isomorphism, Stability, and Reversing Axioms; and let $(a, b; s)$ denote some triple of nonnegative integers in $\mathcal{T}$. It suffices to show that, for all problem instances $A \xrightarrow{f} S \xleftarrow{g} B$ whose dimensions are given by the triple $(a, b; s)$, the quotients $(A \times^*_S B)/(A \times_S B)$ have a common sign, since we can then choose the free bit $w(a, b; s)$ to make that common sign be $(-1)^{w(a,b;s)}$.

Among such problem instances, the generic ones have factor maps f and g that are of the largest possible rank: $\dim(f(A)) = \min(a, s)$ and $\dim(g(B)) = \min(b, s)$. By the Stability Axiom, it suffices for us to show that all such generic instances agree about the sign of the quotient $(A \times^*_S B)/(A \times_S B)$, since any non-generic instance is the limit of a family of generic instances.

Let $A \xrightarrow{f} S \xleftarrow{g} B$ and $A' \xrightarrow{f'} S' \xleftarrow{g'} B'$ be two such generic instances. Since both are generic, we have $\dim(f(A)) = \dim(f'(A'))$ and $\dim(g(B)) = \dim(g'(B'))$. And, since both are transverse, we have $\dim(f(A) + g(B)) = s = \dim(f'(A') + g'(B'))$. It follows that $\dim(f(A) \cap g(B)) = \dim(f'(A') \cap g'(B'))$. Therefore, we can choose a bijection $\chi : S \to S'$ that satisfies $\chi(f(A) \cap g(B)) = f'(A') \cap g'(B')$, $\chi(f(A)) = f'(A')$, and $\chi(g(B)) = g'(B')$. We can then choose bijections $\varphi\colon A \to A'$ and $\psi\colon B \to B'$ that satisfy $\chi(f(\alpha)) = f'(\varphi(\alpha))$ and $\chi(g(\beta)) = g'(\psi(\beta))$.

The three bijections φ, ψ, and χ almost constitute an isomorphism between the unprimed and primed problem instances, as in Diagram 9-1, the only problem being that they might not preserve orientation. So we set $A'' := \pm A'$, $B'' := \pm B'$, and $S'' := \pm S'$, choosing the signs so as to make φ, ψ, and χ preserve orientation as bijections between the unprimed and doubly primed instances. Since both the proper orientation rule and the rule $\times^*$ satisfy the Isomorphism Axiom, the map $(\alpha, \beta) \mapsto (\varphi(\alpha), \psi(\beta))$ will be an orientation-preserving map both from $A \times_S B$ to $A'' \times_{S''} B''$ and also from $A \times^*_S B$ to $A'' \times^*_{S''} B''$. It follows that the two quotients $(A \times^*_S B)/(A \times_S B)$ and $(A'' \times^*_{S''} B'')/(A'' \times_{S''} B'')$ will have the same sign.

Both the proper orientation rule and the rule $\times^*$ satisfy the Left, Right, and Base Reversing Axioms as well, so we have both $A'' \times_{S''} B'' = \eta(A' \times_{S'} B')$ and $A'' \times^*_{S''} B'' = \eta(A' \times^*_{S'} B')$, where the sign $\eta := (A''/A')(B''/B')(S''/S')$ is $+1$ or -1 according as the number of sign reversals between the singly and doubly primed instances is even or odd. It follows that the singly-primed quotient $(A' \times^*_{S'} B')/(A' \times_{S'} B')$ agrees with the other two. $\square$

The Isomorphism and Stability Axioms are our most basic, since they ensure that the orientation rule under discussion will lift from linear spaces to smooth manifolds. The Reversing Axioms are less basic; but it turns out that some appeal to all three of the Reversing Axioms can't be avoided in Proposition 10-1. As the next two exercises show, the Reversing Axioms are needed precisely in those cases where either one of the three input spaces or the output space — that is, either A, B, S, or $A \times_S^* B$ — is zero-dimensional.

Exercise 10-2 In the proof of Proposition 10-1, consider a triple $(a, b; s)$ in which all four of a, b, s, and $a+b-s$ are strictly positive. Show that all problem instances $A \xrightarrow{f} S \xleftarrow{g} B$ with the dimensions $(a, b; s)$ then share a common sign for the quotient $(A[f] \times_S^* [g]B)/(A[f] \times_S [g]B)$ without assuming that the orientation rule $\times^*$ satisfies any of the Reversing Axioms.

Sketch: When $s > 0$, we can choose the bijection $\chi\colon S \to S'$ to preserve orientation. Even when the primed and unprimed problem instances are generic, however, we may then be unable to choose the bijection $\varphi\colon A \to A'$ to preserve orientation, since we might have $\dim(\mathrm{Ker}(f)) = \dim(\mathrm{Ker}(f')) = 0$. But, when a, b, s, and $a + b - s$ are all positive and g has rank $\min(b, s)$, we can afford to lower the rank of the map f from $\min(a, s)$ to $\min(a, s) - 1$ without destroying transversality. Thus, we can smoothly alter the map f to a new map $f^-\colon A \to S$ whose behavior has been negated in precisely one dimension, without losing transversality as the behavior in that dimension passes through zero. The quotient $(A[f^-] \times_S^* [g]B)/(A[f^-] \times_S [g]B)$ for this altered problem instance has the same sign as the original, by stability; and, if we compare that altered instance to the primed instance $A' \xrightarrow{f'} S' \xleftarrow{g'} B'$, using the same bijection χ as before, we can now choose the bijection $\varphi\colon A \to A'$ to preserve orientation. Repeating the same technique, if necessary, we can smoothly alter the map g in one dimension to a new map $g^-\colon B \to S$, thereby allowing the bijection $\psi\colon B \to B'$ to preserve orientation as well — and all without any appeal to the Reversing Axioms.

It follows from this exercise that most instances of the Reversing Axioms are actually consequences of the Isomorphism and Stability Axioms. That is, given the latter two axioms, the only novel content in the Reversing Axioms concerns fiber products where either an input space or the output space is zero-dimensional.

Exercise 10-3 On the other hand, show that the Reversing Axioms are required in Proposition 10-1 for those triples $(a, b; s)$ in which at least one of a, b, s, or $a + b - s$ is zero.

Answer: For triples with $a = 0$, consider the orientation rule $\times^*$ given by

$$A \times_S^* B := \begin{cases} A \times_S B & \text{if } \dim(A) > 0 \\ \Diamond \times_S B & \text{if } \dim(A) = 0. \end{cases}$$

This rule assigns the Proper Orientation except when the factor space $A = -\Diamond$ is both zero-dimensional and negatively oriented, in which case it does the reverse.

So it differs from the Proper Rule by something that isn't simply a dimensional twist. This rule is stable because there is no way to smoothly move from a case that it treats properly to a case that it treats improperly simply by altering the factor maps; changing from $A = \Diamond$ to $A = -\Diamond$ requires a discrete jump. This rule also satisfies the Isomorphism Axiom — and the Right and Base Reversing Axioms, for that matter. But it violates the Left Reversing Axiom.

Triples with $b = 0$ or with $s = 0$ can be handled similarly. For triples with $a+b-s = 0$, consider the orientation rule that orients every positive-dimensional transverse fiber product properly, but orients every zero-dimensional such product positively, whether that is proper or improper.

10.1.2 Which twists have which properties?

For an orientation rule $\times^w$ that is based on a dimensional twist $w\colon \mathcal{T} \to \mathbb{Z}/2\mathbb{Z}$, the structure of the twist w determines which of our remaining axioms will be satisfied. Each axiom boils down to a certain constraint on the twist w. In four cases, the axiom simply forces the twisted rule to agree with the Proper Rule on some subset of $\mathcal{T}$, so the constraint is quite simple:

- The Both Identities Axiom holds when $w(s, s; s) = 0$, for all s in $\mathbb{N}$.
- The Left Identity Axiom holds when $w(s, b; s) = 0$, for all b and s in $\mathbb{N}$.
- The Right Identity Axiom holds when $w(a, s; s) = 0$, for all a and s in $\mathbb{N}$.
- The Concatenate Axiom holds when $w(a, b; 0) = 0$, for all a and b in $\mathbb{N}$.

But the Axiom of Mixed Associativity, $(A \times_S B) \times_T C = A \times_S (B \times_T C)$, is more subtle. If we evaluate the left-hand side using the twisted rule $\times^w$, our result will differ from the proper result by -1 to the power $w(a, b; s)+w(a+b-s, c; t)$. Similarly, the twisted right-hand side will differ from the proper one by -1 to the power $w(a, b+c-t; s) + w(b, c; t)$. So the Axiom of Mixed Associativity will hold just when

$$(10\text{-}4)\qquad w(a, b; s) + w(a + b - s, c; t) = w(a, b + c - t; s) + w(b, c; t),$$

for all nonnegative a, b, c, s, and t with $a+b \geq s$, $b+c \geq t$, and $a+b+c \geq s+t$. We shall refer to Identity 10-4 as the *Mixed Identity*; and we shall refer to those dimensional twists that satisfy the Mixed Identity as being *mixed-associative*.

It is not immediately clear what structural properties of a dimensional twist cause it to be mixed associative. It will turn out that the mixed associative twists are precisely those of the form

$$(10\text{-}5)\qquad w(a, b; s) := u(a) + u(b) + u(a + b - s) + v(s),$$

where $u\colon \mathbb{N} \to \mathbb{Z}/2\mathbb{Z}$ and $v\colon \mathbb{N} \to \mathbb{Z}/2\mathbb{Z}$ are arbitrary functions. Recall that every triple $(a, b; s)$ in $\mathcal{T}$ satisfies $a + b \geq s$, so the argument $a + b - s$ in the third call to u will be nonnegative. We shall refer to Equation 10-5 as the *UV-Recipe*. The twists that are mixed associative, it will turn out, are precisely those that can be constructed from an appropriate pair of functions (u, v) via the UV-Recipe.

If a dimensional twist w can be so constructed from a pair of functions (u, v), then there are always three other pairs of functions (u', v') that would produce the same twist w. First, suppose that we complement all of the values of both functions, setting $u'(k) := u(k) + 1$ and $v'(k) := v(k) + 1$ for all nonnegative k. Since there are four terms on the right-hand side of the UV-Recipe 10-5 and we are working modulo 2, the pair (u', v') will generate the same twist as (u, v). Second, suppose that we complement all of the values $u(k)$ and $v(k)$ for k odd; that is, we set $u'(k) := u(k) + k$ and $v'(k) := v(k) + k$. Since the sum of the four arguments $(a) + (b) + (a + b - s) + (s)$ is zero modulo 2, we again get the same twist. We can eliminate those two sources of redundancy by normalizing the pair (u, v), arbitrarily constraining some two values of the functions u and v, one with an even argument and the other with an odd argument. For example, we could require that $u(1) = u(0) = 0$ or that $v(1) = v(0) = 0$. In fact, we shall constrain one value of u and one value of v, requiring that $u(1) = v(0) = 0$.

Proposition 10-6 *The UV-Recipe 10-5 constitutes a one-to-one correspondence between dimensional twists w that are mixed associative and pairs of functions (u, v) that satisfy the normalization constraints $u(1) = v(0) = 0$.*

Proof It is easy to see that any dimensional twist w that is constructed from a pair of functions (u, v) via the UV-Recipe will satisfy the Mixed Identity 10-4 and will hence be mixed associative. The left-hand side of the Mixed Identity simplifies to

$$\begin{aligned}
&w(a, b; s) + w(a + b - s, c; t)\\
&\quad = u(a) + u(b) + u(a + b - s) + v(s)\\
&\qquad\qquad + u(a + b - s) + u(c) + u(a + b + c - s - t) + v(t)\\
&\quad = u(a) + u(b) + u(c) + u(a + b + c - s - t) + v(s) + v(t),
\end{aligned}$$

and the right-hand side simplifies to that same value.

To show the converse, let w be a twist that satisfies the Mixed Identity 10-4 for all quintuples $(a, b, c; s, t)$ with $a+b \geq s$, $b+c \geq t$, and $a+b+c \geq s+t$. We must construct functions $u\colon \mathbb{N} \to \mathbb{Z}/2\mathbb{Z}$ and $v\colon \mathbb{N} \to \mathbb{Z}/2\mathbb{Z}$ with $u(1) = v(0) = 0$ and with

$$w(a, b; s) = u(a) + u(b) + u(a + b - s) + v(s),$$

for all triples $(a, b; s)$ with $a+b \geq s$. And we must show that this pair of functions (u, v) is uniquely determined.

Step 1: Considering first the quintuples $(s, s, k; s, s)$ for any nonnegative s and k, we find that $w(s, s; s) = w(s, k; s)$. Symmetrically, considering $(k, s, s; s, s)$, we find that $w(s, s; s) = w(k, s; s)$.

Step 2: In the particular case $s = 0$, we have $w(k, 0; 0) = w(0, k; 0) = w(0, 0; 0)$. We want to have $w(0, 0; 0) = u(0)+u(0)+u(0)+v(0) = u(0)+v(0)$, and our normalization requires $v(0) = 0$; so we must set $u(0) := w(0, 0; 0)$. This establishes the UV-Recipe in the case $(0, 0; 0)$.

Step 3: We next consider the quintuples $(a, b, a; 0, 0)$, learning that

$$w(a, b; 0) + w(a + b, a; 0) = w(a, a + b; 0) + w(b, a; 0),$$

which we can rewrite as

$$w(a, a + b; 0) - w(a + b, a; 0) = w(a, b; 0) - w(b, a; 0).$$

From this, an induction that parallels the subtractive algorithm for computing the greatest common divisor allows us to deduce that $w(a, b; 0) = w(b, a; 0)$, for all nonnegative a and b. The base cases of that induction are the cases with $\min(a, b) = 0$, which we established in Step 2.

Step 4: For any $k \geq 1$, we want to arrange that

$$w(1, k; 0) = u(1) + u(k) + u(k + 1) + v(0) = u(k) + u(k + 1),$$

where the latter equality holds because the normalization forces $u(1) = v(0) = 0$. Since we are working modulo 2, it follows that $u(k+1) = u(k)+w(1, k; 0)$. Thus, our only hope is to set $u(k)$, for all $k \geq 1$, as follows: $u(k) := \sum_{1 \leq i < k} w(1, i; 0)$. Note that this gives $u(1) = 0$, as required by the normalization. Furthermore, making this choice for $u(k)$ establishes the UV-Recipe for all of the triples $(a, b; s)$ with $s = 0$ and $\min(a, b) = 1$.

Step 5: Considering next the quintuples $(a, 1, b; 0, 0)$ for any nonnegative a and b, we find that $w(a, 1; 0)+w(a+1, b; 0) = w(a, b+1; 0)+w(1, b; 0)$. Since the UV-Recipe holds for the first and last terms, it holds for the second term just when it holds for the third. By induction along the diagonals in the plane $s = 0$ with a fixed value of $a + b$, we conclude that the UV-Recipe holds for all triples $(a, b; s)$ with $s = 0$.

Step 6: For all $s \geq 0$, we want to arrange that

$$w(s, s; s) = u(s) + u(s) + u(s) + v(s) = u(s) + v(s).$$

Since $u(s)$ is now determined, this forces us to set $v(s) := w(s, s; s) - u(s)$. Note that, when $s = 0$, this gives $v(0) = 0$, as the normalization requires. For $s \geq 1$, this choice establishes the UV-Recipe for the particular triple $(s, s; s)$. Furthermore, we saw in Step 1 that $w(k, s; s) = w(s, k; s) = w(s, s; s)$, for all nonnegative s and k. From this, it follows easily that the UV-Recipe actually holds for all triples of the form $(k, s; s)$ or $(s, k; s)$.

Step 7: Finally, we consider the quintuples $(a, b, 1; s, 0)$ for $a+b \geq s$, finding that $w(a, b; s) + w(a + b - s, 1; 0) = w(a, b + 1; s) + w(b, 1; 0)$. We know that the UV-Recipe holds for the second and fourth terms; so it holds for the first term just when it holds for the third. Thus, by induction, the UV-Recipe holds for the entire column $\{(a, k; s) \mid k \geq \max(0, s - a)\}$ when it holds for any entry in that column. In Step 6, we saw that it does hold for $(a, s; s)$; so it always holds. □

Let's now restrict ourselves to dimensional twists w that are mixed-associative and can therefore be constructed via the UV-Recipe 10-5. For such twists, the Identity and Concatenate Axioms translate into restrictions on the functions u and v. There is an additional, very weak property that is worth considering in this context. It seems natural to hope that $\Diamond \times_{\Diamond} \Diamond = \Diamond$, as opposed to $\Diamond \times_{\Diamond} \Diamond = -\Diamond$. If an orientation rule has the property $\Diamond \times_{\Diamond} \Diamond = \Diamond$, let's say that it *preserves null-positivity*. Note that any rule that satisfies either the Both Identities Axiom or the Concatenate Axiom must preserve null-positivity.

Proposition 10-7 *Recall that every mixed-associative twist corresponds, under the UV-Recipe, to a unique pair of functions* (u, v) *with* $u(1) = v(0) = 0$. *The twists that preserve null-positivity are those with* $u(0) = 0$. *The twists that satisfy the Both Identities Axiom are those with* $u \equiv v$, *that is, with* $u(k) = v(k)$ *for all* $k \geq 0$*; those same twists also satisfy the Left and Right Identity Axioms. The twists that satisfy the Concatenate Axiom are those with* $u \equiv 0$.

Proof A mixed-associative twist w preserves null-positivity just when we have $w(0, 0; 0) = u(0) + u(0) + u(0) + v(0) = u(0) + v(0) = 0$. Since $v(0) = 0$ by assumption, we must also have $u(0) = 0$.

A twist w satisfies the Both Identities Axiom just when we have $w(s, s; s) = u(s) + v(s) = 0$, that is, when $u \equiv v$. When this happens, it also satisfies the Left Identity Axiom, since $w(s, b; s) = u(s) + u(b) + u(b) + v(s) = 0$. And similarly for the Right Identity Axiom.

A twist w satisfies the Concatenate Axiom just when we have $w(a, b; 0) = u(a) + u(b) + u(a + b) + v(0) = u(a) + u(b) + u(a + b) = 0$. Setting $a := k$ and $b := 1$, we find that $u(k) + u(1) + u(k + 1) = 0$; since $u(1) = 0$, this means $u(k) = u(k + 1)$, so we have $u \equiv 0$ by induction. $\square$

Of course, if we want both the Identity and Concatenate Axioms to hold, we must choose $u \equiv v \equiv 0$; that is, we must eschew all twisting and adopt the Proper Orientations.

Exercise 10-8 Let's denote by $\times^{\leftrightarrow}$ the oriented fiber-product operator that would result if we replaced the Concatenate Axiom with the Concatenate-Backwards Axiom, while retaining all of our other axioms. Since all of our other axioms are either left-right symmetric or come in symmetric pairs, it follows that $A \times_S^{\leftrightarrow} B = B \times_S A$, for all binary problem instances $A \xrightarrow{f} S \xleftarrow{g} B$. So the orientation rule $\times^{\leftrightarrow}$ differs from the Proper Rule by the dimensional twist $\leftrightarrow : \mathcal{T} \to \mathbb{Z}/2\mathbb{Z}$ given by $\leftrightarrow (a, b; s) := (a - s)(b - s) \mod 2$. This twist must be mixed-associative; to what functions u and v does it correspond?

Answer: The functions u and v given by

$$u(k) := v(k) := \begin{cases} 0 & \text{if } k \equiv 0, 1 \pmod 4 \\ 1 & \text{if } k \equiv 2, 3 \pmod 4. \end{cases}$$

10.2 The partial identity formulas

Consider a binary problem instance $A \xrightarrow{f} S \xleftarrow{g} B$. If the map f is bijective, then the Left Identity Axiom, together with the Isomorphism and Left Reversing Axioms, tells us at once how to orient the fiber product. Now suppose that f is not bijective, but that we can write f as the direct product of two maps, one of which is bijective. We want to exploit the existence of that bijective factor to simplify the fiber product, and the Left Partial Identity Formula will tell us how.

Changing notation somewhat, let M denote the dimensions that we know to be bijectively mapped, while A and S denote the remainders of the left factor and base spaces. That is, we consider the binary zigzag $(M \times A) \xrightarrow{1 \times f} (M \times S) \xleftarrow{(h,g)} B$. The left factor map is the direct product $1_{M \leftarrow M} \times f$, where $1_{M \leftarrow M} \colon M \to M$ corresponds to the bijectively mapped dimensions and $f \colon A \to S$ to the rest. The right factor map (h, g) also splits into two parts, the map $g \colon B \to S$ and the map $h \colon B \to M$. Is this zigzag transverse? The left factor map certainly covers M, so this zigzag is transverse just when $\mathrm{Im}(f) + \mathrm{Im}(g) = S$, that is, just when the reduced zigzag $A \xrightarrow{f} S \xleftarrow{g} B$ is transverse. Furthermore, when both zigzags are transverse, their fiber products are isomorphic in a canonical way: The map $\Phi \colon A \times B \to M \times A \times B$ given by $\Phi(\alpha, \beta) := (h(\beta), \alpha, \beta)$ carries the reduced fiber product $A \times_S B$ bijectively onto the extended fiber product $(M \times A) \times_{(M \times S)} B$. One might hope that this map Φ would preserve orientation; that is, one might hope that

$$(M \times A) \times_{(M \times S)} B \stackrel{?}{=} A \times_S B. \tag{10-9}$$

If so, this would be exactly the type of formula that we were looking for: Given a fiber product in which the left factor map is a direct product with a bijective factor, we could simplify by removing the domain and codomain of that bijection.

Unfortunately, Formula 10-9 turns out to be wrong; more precisely, it turns out to be inconsistent with the Concatenate Axiom. The problem is that the factor of M multiplies A and S from the left. We set things up that way without discussing the issue. But we can equally well put the bijectively mapped dimensions on the right, forming the extended zigzag $(A \times M) \xrightarrow{f \times 1} (S \times M) \xleftarrow{(g,h)} B$. This suggests an analogous formula in which M multiplies A and S from the right:

$$(A \times M) \times_{(S \times M)} B = A \times_S B. \tag{10-10}$$

Formulas 10-9 and 10-10 can't both be right, since they give different answers in some cases — for example, when $\dim(A)$ and $\dim(M)$ are both odd, so that $A \times M = -(M \times A)$, but $\dim(S)$ is even, so that $S \times M = M \times S$. The special cases in which $S = \lozenge$ are particularly revealing. In those cases, we shall see that Formula 10-10 follows easily from our axioms. Since Formula 10-9 disagrees with Formula 10-10 on some of those cases, Formula 10-9 is definitely wrong.

When $S = \Diamond$, Formula 10-10 claims that $(A \times M) \times_{(\Diamond \times M)} B = A \times_\Diamond B$. This case is particularly simple to analyze because the understood left factor map $f\colon A \to S$ must be identically zero. Applying the Concatenate Axiom to the direct product $A \times M$, we have

$$(A \times M) \times_{(\Diamond \times M)} B = (A \times_\Diamond M) \times_{(\Diamond \times M)} B.$$

Since $f \equiv 0$, the left factor map of the outer fiber product on the right-hand side is the composite map $0 \times 1\colon A \times_\Diamond M \to \Diamond \times M$. The value of this map at a point (α, μ) doesn't depend upon α; so the side condition discussed in Section 9.1.7 is met. Thus, we can apply the Axiom of Mixed Associativity to the transverse, ternary zigzag

$$A \xrightarrow{0} \Diamond \xleftarrow{0} M \xrightarrow{(0,1)} \Diamond \times M \xleftarrow{(0,g)} B,$$

to swap the parentheses from the left to the right:

$$(A \times_\Diamond M) \times_{(\Diamond \times M)} B = A \times_\Diamond (M \times_{(\Diamond \times M)} B) = A \times_\Diamond (M \times_M B).$$

The Left Identity Axiom then reduces this answer to $A \times_\Diamond B$, which establishes Formula 10-10 in the case $S = \Diamond$.

The upcoming Proposition 10-12 will show that all cases of Formula 10-10 follow from our axioms; so we shall henceforth refer to Formula 10-10 as the *Left Partial Identity Formula*. Proposition 10-12 will also establish the symmetric formula for those binary fiber products in which the right factor map is partially bijective, which is the *Right Partial Identity Formula*

$$A \times_{(M \times S)} (M \times B) = A \times_S B. \tag{10-11}$$

Proposition 10-12 *The zigzag* $(A \times M) \xrightarrow{f \times 1} (S \times M) \xleftarrow{(g,h)} B$ *is transverse just when the reduced zigzag* $A \xrightarrow{f} S \xleftarrow{g} B$ *is transverse, and, under the Proper Orientations, we have* $(A \times M) \times_{(S \times M)} B = A \times_S B$. *Symmetrically, the zigzag* $A \xrightarrow{(e,f)} (M \times S) \xleftarrow{1 \times g} (M \times B)$ *is transverse just when* $A \xrightarrow{f} S \xleftarrow{g} B$ *is transverse, and we have* $A \times_{(M \times S)} (M \times B) = A \times_S B$.

Proof The claims about transversality hold because the identity map $1\colon M \to M$ covers M, leaving $\mathrm{Im}(f)$ and $\mathrm{Im}(g)$ to cover S.

As for the Partial Identity Formulas themselves, it suffices to verify one of them, since each implies the other. For example, if we assume the Right Partial Identity Formula 10-11, we can then prove the Left Formula 10-10 by keeping track of the sign changes as we commute the various direct and fiber products:

$$\begin{aligned}
(A \times M) \times_{(S \times M)} B &= (-1)^{am+sm}\big((M \times A) \times_{(M \times S)} B\big)\\
&= (-1)^{am+sm+((m+a)-(m+s))(b-(m+s))}\big(B \times_{(M \times S)} (M \times A)\big)\\
&= (-1)^{am+sm+(a-s)(b-m-s)}\big(B \times_S A\big)\\
&= (-1)^{am+sm+(a-s)(b-m-s)+(b-s)(a-s)}\big(A \times_S B\big)\\
&= A \times_S B.
\end{aligned}$$

To verify the Right Partial Identity Formula, it suffices to show that it holds under the Calibrated Delta Rule, since that rule assigns the Proper Orientations. So we consider the equation

$$A[(e,\ f)] \times_{(M\times S)} [1 \times g](M \times B) = A[f] \times_S [g]B.$$

Let $P := A \times_S B$ denote the right-hand fiber product, viewed as a subset of the direct product $A \times B$, and let $\Phi\colon A \times B \to A \times M \times B$ be the map given by $\Phi(\alpha, \beta) := (\alpha, e(\alpha), \beta)$. The left-hand fiber product $P' := A \times_{(M\times S)} (M \times B)$ then satisfies $P' = \eta\,\Phi(P)$ for some sign $\eta = \pm 1$, and we claim that $\eta = +1$. Let C be any oriented complement of P in $A \times B$. Applying the Delta Rule to the right-hand fiber product, we have

$$\frac{P \oplus C}{A \times B} = (-1)^{s(b-s)} \frac{\Delta(C)}{S}. \tag{10-13}$$

Turning to the left-hand fiber product, we choose $C' := \Phi(C) \oplus (0, M, 0)$ as our oriented complement of P' in $A \times M \times B$. The Delta Rule then gives us

$$\frac{P' \oplus C'}{A \times M \times B} = (-1)^{(m+s)((m+b)-(m+s))} \frac{\Delta'(C')}{M \times S}. \tag{10-14}$$

Rewriting P' as $\eta\,\Phi(P)$ and substituting in for C', we deduce that

$$\eta\,\frac{\Phi(P) \oplus \Phi(C) \oplus (0, M, 0)}{A \times M \times B} = (-1)^{(m+s)(b-s)} \frac{\Delta'(\Phi(C) \oplus (0, M, 0))}{M \times S}. \tag{10-15}$$

We now transform the left-hand side of Equation 10-13 so as to make it to resemble the left-hand side of Equation 10-15. Since the map Φ is injective, we can apply it to both numerator and denominator, getting $(P \oplus C)/(A \times B) = \big(\Phi(P) \oplus \Phi(C)\big)/\Phi(A \times B)$. Since $\Phi(A \times B)$ and $(0, M, 0)$ are linearly independent subspaces of $A \times M \times B$, we can then add a direct summand of $(0, M, 0)$ to both numerator and denominator, getting

$$\frac{\Phi(P) \oplus \Phi(C) \oplus (0, M, 0)}{\Phi(A \times B) \oplus (0, M, 0)} = (-1)^{s(b-s)} \frac{\Delta(C)}{S}. \tag{10-16}$$

The numerators on the left-hand sides of Equations 10-15 and 10-16 are the same. As for the denominator in Equation 10-16, we have $\Phi(A\times B)\oplus(0, M, 0) = \Psi(A \times B \times M)$, where $\Psi\colon A \times B \times M \to A \times M \times B$ is the linear map with the matrix

$$\Psi = \begin{pmatrix} 1_{A\leftarrow A} & 0_{A\leftarrow B} & 0_{A\leftarrow M} \\ e & 0_{M\leftarrow B} & 1_{M\leftarrow M} \\ 0_{B\leftarrow A} & 1_{B\leftarrow B} & 0_{B\leftarrow M} \end{pmatrix}.$$

The entry of e here can be eliminated by an elementary operation on either the rows or the columns, so we have $\Psi(A \times B \times M) = (-1)^{bm}(A \times M \times B)$. Combining this result with Equations 10-15 and 10-16, we deduce that

$$\eta\,(-1)^{bm+s(b-s)} \frac{\Delta(C)}{S} = (-1)^{(s+m)(b-s)} \frac{\Delta'(\Phi(C) \oplus (0, M, 0))}{M \times S}. \tag{10-17}$$

Now, since $\Delta'(\Phi(\alpha,\beta)) = \Delta'(\alpha, e(\alpha), \beta) = (e(\alpha), g(\beta)) - (e(\alpha), f(\alpha)) = (0, g(\beta) - f(\alpha))$, we have $\Delta'(\Phi(C)) = (0, \Delta(C))$. And since $\Delta'(0,\mu,0) = (\mu, 0) - (0,0) = (\mu, 0)$, we have $\Delta'(0, M, 0) = (M, 0)$. We deduce that

$$\eta\,\frac{\Delta(C)}{S} = (-1)^{ms}\frac{(0,\Delta(C)) \oplus (M,0)}{(M,0)\oplus(0,S)} = \frac{(M,0)\oplus(0,\Delta(C))}{(M,0)\oplus(0,S)} = \frac{M\times\Delta(C)}{M\times S},$$

and so the sign η is indeed $+1$, as we claimed. $\square$

Exercise 10-18 Suppose that we chose to replace the Concatenate Axiom with the Concatenate-Backwards Axiom. Show that the Partial Identity Formulas above would then be replaced by the *Left Partial Identity Back-Formula*

$$(M \times A) \times^{\longleftrightarrow}_{(M\times S)} B = A \times^{\leftrightarrow}_S B \tag{10-19}$$

and the *Right Partial Identity Back-Formula*

$$A \times^{\longleftrightarrow}_{(S\times M)} (B \times M) = A \times^{\leftrightarrow}_S B, \tag{10-20}$$

where $\times^{\leftrightarrow}$ is the Concatenate-Backwards version of the oriented fiber product that we introduced in Exercise 10-8. The former of these can be viewed as a corrected version of the erroneous Formula 10-9, with which we began Section 10.2.

Answer: If we swap the left and right factor spaces on both sides of the Left Partial Identity Formula 10-10, we get $B \times^{\longleftrightarrow}_{(S\times M)} (A \times M) = B \times^{\leftrightarrow}_S A$. This becomes the Right Partial Identity Back-Formula 10-20 when we restore alphabetical order by swapping A and B. The Left Partial Identity Back-Formula 10-19 follows in a similar way from the Right Partial Identity Formula 10-11.

Exercise 10-21 For n a positive integer, suppose that the oriented linear space $S = A_1 \times \cdots \times A_n$ is the direct product of n oriented spaces. Show that

$$(\hat{A}_1 \times \cdots \times A_n) \times_S (A_1 \times \hat{A}_2 \times \cdots \times A_n) \times_S \cdots \times_S (A_1 \times \cdots \times \hat{A}_n) = \lozenge,$$

where a hat indicates a factor that is omitted from a direct product and where the factor maps of the fiber product on the left-hand side are the obvious inclusions. Note that, when $n = 2$, this reduces to the identity $B \times_{(A\times B)} A = \lozenge$ that we discussed in Exercise 9-8.

Answer: Consider the k^{th} factor space $A_1 \times \cdots \times A_{k-1} \times A_{k+1} \times \cdots \times A_n$, for some k in $[1 \,..\, n]$. For i from n down to $k+1$, working from right to left, we use the Left Partial Identity Formula to cancel the factor of A_i in this product against the A_i in the next base space S to the right. Similarly, for i from 1 to $k-1$, working from left to right, we use the Right Partial Identity Formula to cancel the A_i in this product against the A_i in the next S to the left. All of this canceling leaves $\lozenge \times_\lozenge \cdots \times_\lozenge \lozenge$, which is simply $\lozenge$.

10.3 The Binary Full Formula

The Partial Identity Formulas are less transparent than our individual axioms, but they are also more powerful. Moving even further in that direction, we can combine both of the Partial Identity Formulas with the Concatenate Axiom into a single, subtle formula, the *Binary Full Formula*:

$$(L \times N \times Q) \times_{(M \times N \times Q)} (M \times N \times R) = L \times N \times R. \tag{10-22}$$

If λ, μ, ν, θ, and ρ are elements of the spaces L, M, N, Q, and R, then the left-hand factor map here takes $(\lambda, \nu, \theta) \mapsto (0, \nu, \theta)$, while the right-hand map takes $(\mu, \nu, \rho) \mapsto (\mu, \nu, 0)$. That is, we are dealing with the binary problem instance

$$L \times N \times Q \xrightarrow{0\times 1\times 1} M \times N \times Q \xleftarrow{1\times 1\times 0} M \times N \times R.$$

The letters L, M, N, Q, and R are less than perspicuous. A better way to name the five oriented linear spaces in the Binary Full Formula is to give each space a two-letter name, where those two letters are drawn, with repetition, from the set $\{A, B, S\}$. Under this scheme, the Binary Full Formula is written

$$(AA \times AB \times AS) \times_{(SB \times AB \times AS)} (SB \times AB \times BB) = AA \times AB \times BB. \tag{10-23}$$

To start to get a handle on this formula, let's consider a transverse fiber product $P := A[f] \times_S [g]B$ and think about the ways in which the various dimensions of the linear spaces A, S, and B can interact.

AB: There may be some dimensions of A, B, and S that are in mutual bijective correspondence under f and g. In the Binary Full Formula 10-23, those dimensions make up AB, and they survive into the fiber product P.

AS: There may be some dimensions of A that are in bijective correspondence with some dimensions of S under f, but where those dimensions of S lie outside the image of g. Such dimensions make up AS. They don't survive into the fiber product P — because Bob couldn't stay at the same altitude as Alice if Alice moved in those dimensions.

SB: Symmetrically, SB consists of those dimensions of B that are in bijective correspondence with dimensions of S under g, but where those dimensions of S lie outside the image of f.

AA: There may be some dimensions of A that lie in the kernel of f — ways that Alice can move without affecting her altitude. Those dimensions make up AA, and they survive into the fiber product P.

BB: Symmetrically, BB is those dimensions of B that lie in the kernel of g.

Note that the two letters in the name of a space should be thought of as the left and right endpoints of a nonempty substring of the string "*ASB*". In particular, the space AB appears in the S position of Formula 10-23, as well as in the A and B positions. Note also that we are using only five out of the six possible such substrings. By transversality, every dimension of S must correspond either to a dimension of A (Case AS) or a dimension of B (Case SB) or both (Case AB). Thus, we have no need for a "neither" option, a space named SS.

The Binary Full Formula 10-23 follows easily from the two Partial Identity Formulas and the Concatenate Axiom. Starting with the left-hand side

$$(AA \times AB \times AS) \times_{(SB \times AB \times AS)} (SB \times AB \times BB),$$

we apply the Left Partial Identity Formula to remove the two instances of AS and the Right Partial Identity Formula to remove the two instances of SB, leaving $(AA \times AB) \times_{AB} (AB \times BB)$. We then choose one of the two Partial Identity Formulas — say the Left one — to remove two of the three instances of AB, after which the Concatenate Axiom finishes the proof: $(AA \times AB) \times_{AB} (AB \times BB) = AA \times_{\Diamond} (AB \times BB) = AA \times (AB \times BB) = AA \times AB \times BB$.

While the Binary Full Formula is easy to prove, it is surprisingly powerful. The key to that power — the reason why it is "full" — is the following universality.

Proposition 10-24 *Every binary problem instance* $A \xrightarrow{f} S \xleftarrow{g} B$ *is isomorphic to an instance of the special form*

$$AA \times AB \times AS \xrightarrow{0\times 1\times 1} SB \times AB \times AS \xleftarrow{1\times 1\times 0} SB \times AB \times BB,$$

to which the Binary Full Formula applies.

Proof We want to express the oriented factor spaces A and B as direct products $A = AA \times AB \times AS$ and $B = SB \times AB \times BB$ and the base space S as a direct product $S = SB \times AB \times AS$ in such a way that the factor maps f and g become the trivial direct products $0 \times 1 \times 1$ and $1 \times 1 \times 0$. To this end, we shall construct oriented subspaces A_{AA}, A_{AB}, and A_{AS} of A with $A = A_{AA} \oplus A_{AB} \oplus A_{AS}$ and similarly for B and S, where the factor maps f and g restrict to give orientation-preserving bijections relating $A_{AS} \leftrightarrow S_{AS}$, $S_{SB} \leftrightarrow B_{SB}$, and $A_{AB} \leftrightarrow S_{AB} \leftrightarrow B_{AB}$.

It's convenient to do the work in two phases. We'll first do linear algebra to choose the subspaces themselves, but just orient them arbitrarily. We'll then go back and reverse the orientations of some of those chosen subspaces, as necessary, so that the final orientations of A, B, and S match the given problem instance.

The linear algebra starts out easily; we set $A_{AA} := \mathrm{Ker}(f)$ and $B_{BB} := \mathrm{Ker}(g)$, assigning orientations to those two kernels arbitrarily.

We next fix some complement U of A_{AA} in A. If we restrict the factor map f to the subspace U of A, it becomes a bijection onto $\mathrm{Im}(f) \subseteq S$. Let's refer to that restriction as the map $\bar{f}\colon U \to \mathrm{Im}(f)$. In a similar way, fix some complement V of B_{BB} in B, and let $\bar{g}\colon V \to \mathrm{Im}(g)$ be the resulting bijection.

We next set $S_{AB} := \mathrm{Im}(f) \cap \mathrm{Im}(g)$, once again orienting arbitrarily. We then set $A_{AB} := \bar{f}^{-1}(S_{AB})$, so that A_{AB} is some oriented linear subspace of U, our chosen complement of A_{AA} in A. Similarly, we set $B_{AB} := \bar{g}^{-1}(S_{AB})$.

The substrings that remain are AS and SB. We set S_{AS} to be some complement of S_{AB} in $\mathrm{Im}(f)$, oriented arbitrarily. We then set $A_{AS} := \bar{f}^{-1}(S_{AS})$. Similarly, we set S_{SB} to be some complement of S_{AB} in $\mathrm{Im}(g)$, oriented arbitrarily, and we set $B_{SB} := \bar{g}^{-1}(S_{SB})$.

We now have $S_{AB} \oplus S_{AS} = \pm \mathrm{Im}(f)$. Applying the bijection $\bar{f}^{-1}$ to each summand, we have $A_{AB} \oplus A_{AS} = \pm U$. And it follows that $A_{AA} \oplus A_{AB} \oplus A_{AS} = \pm(A_{AA} \oplus U) = \pm A$. Similarly, we have $S_{SB} \oplus S_{AB} = \pm \mathrm{Im}(g)$ and, applying $\bar{g}^{-1}$, we have $B_{SB} \oplus B_{AB} = \pm V$ and $B_{SB} \oplus B_{AB} \oplus B_{BB} = \pm B$. Furthermore, since the maps f and g are transverse, their images span all of S; so we have $S_{SB} \oplus S_{AB} \oplus S_{AS} = \pm S$.

We now fix up the orientations, as necessary, starting with the base space S. If the orientation of the direct sum $S_{SB} \oplus S_{AB} \oplus S_{AS}$ is currently $-S$, we reverse the orientation of, say, the summands S_{AS} and A_{AS}. Since we are reversing both of them, the map f still restricts to an orientation-preserving bijection between them; but we now have $S_{SB} \oplus S_{AB} \oplus S_{AS} = +S$. We then fix up the factor spaces. If the orientation of the direct sum $A_{AA} \oplus A_{AB} \oplus A_{AS}$ is currently $-A$, we simply reverse the orientation of the summand A_{AA}. Similarly, if the orientation of $B_{SB} \oplus B_{AB} \oplus B_{BB}$ is currently $-B$, we reverse B_{BB}. The resulting problem instance is then isomorphic to the original. □

Corollary 10-25 *The Binary Full Formula 10-23 and the Isomorphism Axiom — just those two, without any of our other axioms, not even Stability! — form a consistent and complete axiom system for the Proper Orientations.*

Proof Given any binary problem instance $A \xrightarrow{f} S \xleftarrow{g} B$, Proposition 10-24 constructs an isomorphic instance of the special form to which the Binary Full Formula applies. The Proper Orientation for the fiber product $P := A \times_S B$ is then determined as the orientation that makes $P = AA \times AB \times BB$. □

So the Binary Full Formula, together with the Isomorphism Axiom, implies all of our other axioms; but our proof of that in Corollary 10-25 is based on much of this monograph. It is instructive to confirm that the Binary Full Formula implies each of our other axioms, using separate arguments that are as direct as possible. We are going to devote the rest of Section 10.3 to that task, since those confirming arguments help to explain just how the Binary Full Formula manages to bundle up so much information into a single equation.

By the way, analogous to our Binary Full Formula, there exists an n-ary Full Formula for any n. The Unary Full Formula is trivial, but is worth writing down, to help clarify the general patterns: $AA = AA$. We'll write out the Ternary Full Formula in Section 10.3.4, and we'll tackle the n-ary case in Section 10.4.

10.3.1 The easy confirmations

The Concatenate and Both Identities Axioms are the easiest axioms to confirm as being implied by the Binary Full Formula — so easy that we don't even need the Isomorphism Axiom. Setting $AA := A$ and $BB := B$ in the Binary Full Formula while setting $AS := AB := SB := \Diamond$, we find that

$$(A \times \Diamond \times \Diamond) \times_{(\Diamond \times \Diamond \times \Diamond)} (\Diamond \times \Diamond \times B) = A \times \Diamond \times B,$$

from which it follows immediately that $A \times_{\Diamond} B = A \times B$. And setting $AB := S$ while setting $AA := AS := SB := BB := \Diamond$, we find that

$$(\Diamond \times S \times \Diamond) \times_{(\Diamond \times S \times \Diamond)} (\Diamond \times S \times \Diamond) = \Diamond \times S \times \Diamond,$$

so we have confirmed that $S \times_S S = S$.

The Left Identity Axiom is a bit more subtle. If we set $AA := SB := \Diamond$ in the Binary Full Formula, we find that

$$(\Diamond \times AB \times AS) \times_{(\Diamond \times AB \times AS)} (\Diamond \times AB \times BB) = \Diamond \times AB \times BB.$$

This is perfectly consistent with the axiom $S \times_S B = B$, where the role of S is played by $AB \times AS$ and the role of B by $AB \times BB$; but we don't yet have a proof. Indeed, in the Left Identity Axiom $S \times_S B = B$, we don't need to assume anything about the behavior of the right factor map $g\colon B \to S$. In order to apply the Binary Full Formula, however, we must break up the spaces B and S into subspaces on which g behaves in known, simple ways. Proposition 10-24 tells us that this breaking up can always be done, up to isomorphism. So the Binary Full Formula, together with the Isomorphism Axiom, does imply the Left Identity Axiom. The Right Identity Axiom is symmetric.

The Reversing Axioms are a similar story. Every binary problem instance $A \times_S B = P$ is isomorphic to an instance of the special form to which the Binary Full Formula applies:

$$(AA \times AB \times AS) \times_{(SB \times AB \times AS)} (SB \times AB \times BB) = AA \times AB \times BB.$$

By negating AA, we then conclude that $(-A) \times_S B = -(A \times_S B)$; and negating BB gives us the Right Reversing Axiom in a similar way. For the Base Reversing Axiom, we negate, say, both AS and AA.

10.3.2 Confirming stability

The Stability Axiom is so subtle that we shan't bother to confirm it here in detail. But we shall discuss the key idea that would underlie such a confirmation.

The hard part of confirming stability is showing that the orientation of the fiber product doesn't suddenly flip at the end of a limiting process in which the rank of a factor map drops. Suppose that we approach a particular binary problem instance

$A \xrightarrow{f_0} S \xleftarrow{g_0} B$ along a path of instances $A \xrightarrow{f_t} S \xleftarrow{g_t} B$, as t tends to zero. We always have $\operatorname{rank}(f_0) \leq \lim_{t\to 0} \operatorname{rank}(f_t)$ and $\operatorname{rank}(g_0) \leq \lim_{t\to 0} \operatorname{rank}(g_t)$. When both of those hold as equalities, stability actually follows from the Isomorphism Axiom. The hard cases are those in which at least one of the inequalities is strict.

Let's suppose that $\operatorname{rank}(f_0) < \lim_{t\to 0} \operatorname{rank}(f_t)$; what happens when we apply the Binary Full Formula to each problem instance $A[f_t] \times_S [g_t]B$? In particular, how does Proposition 10-24 decompose the spaces A, B, and S into subspaces, as a function of t? And how does the Binary Full Formula then build up the fiber product P from those subspaces? At the moment that t becomes zero, some dimensions move out of the subspace AB. In the factor space A, those dimensions move out of A_{AB} into A_{AA}; that is, they get absorbed into the kernel of the map f_t, which suddenly grows. In the fiber product P, the corresponding dimensions move out of P_{AB} into P_{AA}. The corresponding dimensions of the base space S move out of S_{AB} into S_{SB}; that is, they are left behind by the image of f_t, which suddenly shrinks. Note that, since these dimensions of S are no longer covered by $\operatorname{Im}(f_t)$ when $t = 0$, they must be covered by $\operatorname{Im}(g_0)$, since the problem instance $A[f_0] \times_S [g_0]B$ is transverse. So they do indeed move into S_{SB}, as we claimed. The corresponding dimensions in the factor space B move out of B_{AB} into B_{SB}.

We now exploit a combinatorial property of the Binary Full Formula. Note that its four ternary direct products correspond precisely to the four left-to-right chains in the following poset:

(10-26)
$$\begin{tikzcd}[column sep=small, row sep=small]
AA \arrow[dr, dash] & & BB \\
& AB \arrow[ur, dash] \arrow[dr, dash] & \\
SB \arrow[ur, dash] & & AS
\end{tikzcd}$$

In particular, the subspaces AA and SB appear only in direct products with AB immediately to their right, while AB always has either AA or SB to its left. We can thus proceed as follows: As t approaches zero, we arrange that the "moving" dimensions come first, in our positive, ordered bases for each of the four subspaces A_{AB}, S_{AB}, B_{AB}, and P_{AB}. At the moment that t becomes zero, those dimensions can then slide over, with no transpositions needed, to become the last dimensions in our positive, ordered bases for A_{AA}, S_{SB}, B_{SB}, and P_{AA}. Since the order of the basis elements doesn't change as t becomes zero — all that changes is our interpretation of which basis elements span which subspaces — the orientations that are produced by the Binary Full Formula do remain stable in the limit.

We could deal in a similar way with a sudden drop in the rank of g_t when t becomes zero, because the subspaces AS and BB appear always just to the right of AB. We could even deal with the ranks of both f_t and g_t dropping when t becomes zero, since the drop in $\operatorname{rank}(f_t)$ shrinks the basis for AB from the front, while the drop in $\operatorname{rank}(g_t)$ shrinks it from the back — so the two processes don't interfere. Thus, if we filled in lots of details, we could confirm that the Binary Full Formula (together with the Isomorphism Axiom) implies the Stability Axiom.

10.3.3 Reordering the direct products

Writing the Binary Full Formula 10-23 as we have been doing,

$$(AA \times AB \times AS) \times_{(SB \times AB \times AS)} (SB \times AB \times BB) = AA \times AB \times BB,$$

with its direct products in the order specified by Diagram 10-26, works well for confirming stability. But it is worth noting that there are other ways to order those products that are equally valid.

For example, the subspace SB appears twice, with AB to its right both times. If we swapped that pair in one of the two products, putting SB to the right of AB, we would introduce a correction factor of $(-1)^{\dim(SB)\dim(AB)}$. But swapping the other pair as well removes the correction factor, giving us an equally valid way to write the Binary Full Formula:

$$(AA \times AB \times AS) \times_{(AB \times SB \times AS)} (AB \times SB \times BB) = AA \times AB \times BB.$$

The ternary direct products in this variant are left-to-right chains in this poset:

(10-27)
$$\begin{tikzcd}[row sep=tiny] & & & BB \\ AA \arrow[r, dash] & AB \arrow[r, dash] & SB \arrow[ur, dash] \arrow[dr, dash] & \\ & & & AS \end{tikzcd}$$

This variant of the Binary Full Formula makes it harder to write down the left factor map f, however. We can't just write $f = 0 \times 1 \times 1$, because corresponding subspaces no longer occur in corresponding left-to-right positions. Instead, if we wrote $f \colon AA \times AB \times AS \to AB \times SB \times AS$ as a matrix of maps, it would be

$$f = \begin{pmatrix} 0_{AB \leftarrow AA} & 1_{AB \leftarrow AB} & 0_{AB \leftarrow AS} \\ 0_{SB \leftarrow AA} & 0_{SB \leftarrow AB} & 0_{SB \leftarrow AS} \\ 0_{AS \leftarrow AA} & 0_{AS \leftarrow AB} & 1_{AS \leftarrow AS} \end{pmatrix},$$

with not all of the identity blocks on the main diagonal. But the factor map that we intend is still clear: The matrix of maps should have identity blocks wherever the domain and codomain correspond and should otherwise be entirely zero. This rule makes perfect sense even when the matrix of maps is not square, as will happen in constructing the factor maps for the Ternary and n-ary Full Formulas.

Here is another possible variant of the Binary Full Formula: After swapping SB with AB, we could swap AA with AB, both times that that pair appears, getting the variant

$$(AB \times AA \times AS) \times_{(AB \times SB \times AS)} (AB \times SB \times BB) = AB \times AA \times BB,$$

with its poset:

(10-28)
$$\begin{tikzcd}[row sep=small] & AA \arrow[r, dash] \arrow[dr, dash] & BB \\ AB \arrow[ur, dash] \arrow[dr, dash] & & \\ & SB \arrow[r, dash] \arrow[ur, dash] & AS \end{tikzcd}$$

In this variant, the AB component comes first in all four of the direct products.

On the other hand, there certainly are constraints on the posets that we can adopt. For example, we must have $AA \prec BB$ (that is, AA to the left of BB) in any valid poset. To see why, consider a problem instance in which AA and BB are odd-dimensional, while $AB = AS = SB = \Diamond$. The left-hand side of the Binary Full Formula then reduces to $AA \times_{\Diamond} BB$, while the right-hand side is either $AA \times BB$ or $BB \times AA$, according as $AA \prec BB$ or $BB \prec AA$ in the poset. In order to satisfy the Concatenate Axiom, the former must pertain.

Exercise 10-29 Show, by a similar argument based on the Left Partial Identity Formula, that $AA \prec AS$ in any valid poset. Symmetrically, because of the Right Partial Identity Formula, we must have $SB \prec BB$.

Exercise 10-30 Show, by using both of the partial identity formulas (or by using the formula $B \times_{A \times B} A = \Diamond$ in Exercise 9-8) that $SB \prec AS$ in any valid poset.

The four constraints $AA \prec BB$, $AA \prec AS$, $SB \prec BB$, and $SB \prec AS$, which are the four in Poset 10-28 that don't involve AB, turn out to be the only ones that must hold in any valid poset for the Binary Full Formula. We demonstrate that, and generalize to the n-ary case, in Section 10.4.2.

10.3.4 Confirming mixed associativity

We have left the most intriguing axiom for last: Why is it that the Binary Full Formula (together with the Isomorphism Axiom) implies the Axiom of Mixed Associativity? The obvious strategy for confirming that implication starts with a ternary problem instance $A \xrightarrow{f} S \xleftarrow{g} B \xrightarrow{h} T \xleftarrow{k} C$. We then confirm that

$$\Big(A[f] \times_S [g]B\Big)[h'] \times_T [k]C = A[f] \times_S [g']\Big(B[h] \times_T [k]C\Big)$$

by applying the Binary Full Formula four times to evaluate the inner and outer fiber products on each side. In order to apply the Binary Full Formula, however, we have to decompose the various linear spaces involved into subspaces that the various factor maps behave on in known, simple ways. Proposition 10-24 does that decomposition in the binary case, breaking up each of the three spaces A, S, and B into subspaces corresponding to the appropriate subset of the five labels AA, SB, AB, AS, and BB. In the ternary case, the decomposition is more complex. We must break up the five spaces A, S, B, T, and C, and we have a total of twelve potential labels: all fifteen non-empty substrings of the string "$ASBTC$",

$$\begin{array}{c} AA \quad SS \quad BB \quad TT \quad CC \\ AS \quad SB \quad BT \quad TC \\ AB \quad ST \quad BC \\ AT \quad SC \\ AC, \end{array} \tag{10-31}$$

except for SS, ST, and TT, which are ruled out by transversality. Rather than tackling that ternary decomposition here, we postpone it until Section 10.4.1, in which we prove the n-ary analog of Proposition 10-24 by appealing to the theory of quiver representations. For now, let's simply assume the ternary analog.

Under that assumption, the following is a universal formula for the ternary case, which we therefore christen our *Ternary Full Formula*:

$$\mathcal{A} \times_{\mathcal{S}} \mathcal{B} \times_{\mathcal{T}} \mathcal{C} = \mathcal{P} \tag{10-32}$$

where

$$\begin{aligned}
\mathcal{A} &:= AA \times AB \times AC \times AT \times AS \\
\mathcal{S} &:= SB \times SC \times AB \times AC \times AT \times AS \\
\mathcal{B} &:= SB \times SC \times AB \times AC \times BB \times BC \times AT \times BT \\
\mathcal{T} &:= TC \times SC \times AC \times BC \times AT \times BT \\
\mathcal{C} &:= TC \times SC \times AC \times BC \times CC \\
\mathcal{P} &:= AA \times AB \times AC \times BB \times BC \times CC
\end{aligned}$$

This Ternary Full Formula is, unfortunately, far too long to write on a single line; so we have introduced symbolic names for each of its six direct products.

Let's consider first which subspaces appear in which direct products, without worrying about their order. The subspaces that appear in the five products on the left-hand side are determined by the order of the letters in the string "*ASBTC*". For example, for the product $\mathcal{T}$, we start with the eight substrings that contain the letter T — eight because there are four gaps to the left of T, where the substring can start, and two gaps after T, where it can stop. These eight substrings form the parallelogram that can be reached by downward paths from TT in Diagram 10-31. But we then omit the two spaces ST and TT, which are ruled out be transversality, leaving $\mathcal{T}$ as the product of six subspaces. The fiber product $\mathcal{P}$ is a different story; it contains all six of the subspaces whose names involve neither S nor T.

The order of the subspaces in each direct product is a subtler issue. As we have chosen to write down the Ternary Full Formula 10-32, its six direct products correspond to chains in this poset:

$$\begin{tikzcd}
 & AA \arrow[dr, dash] & & & & & CC & \\
TC \arrow[dr, dash] & & AB \arrow[r, dash] & AC \arrow[r, dash] & BB \arrow[r, dash] & BC \arrow[ur, dash] \arrow[dr, dash] & & AS \\
 & SC \arrow[ur, dash] & & & & & AT \arrow[ur, dash] \arrow[dr, dash] & \\
SB \arrow[ur, dash] & & & & & & & BT
\end{tikzcd} \tag{10-33}$$

As in the binary case, there are lots of other orderings that would be equally good. Note that we haven't been able to achieve perfect left-right symmetry in the ternary case. The subspaces AC and BB appear together in several products, so we had to break the symmetry; we have chosen arbitrarily to write AC to the

$$
\begin{aligned}
\mathcal{A} \times_{\mathcal{S}} \mathcal{B} &= (AA, AB, AC, AT, AS) \times_{(SB,SC,AB,AC,AT,AS)} (SB, SC, AB, AC, BB, BC, AT, BT) \\
(-1)^{(at)(bb+bc)} \mathcal{A} \times_{\mathcal{S}} \mathcal{B} &= (AA, \underbrace{AB, AC, AT}, AS) \times_{(\underbrace{SB,SC},\underbrace{AB,AC,AT},AS)} (\underbrace{SB, SC}, \underbrace{AB, AC, AT}, \underbrace{BB, BC, BT}) \\
(-1)^{(at)(bb+bc)} \mathcal{L} &= (AA, \underbrace{AB, AC, AT}, \underbrace{BB, BC, BT}) \\
\mathcal{L} &= (AA, AB, AC, BB, BC, AT, BT) \\
(-1)^{(bb)(ac)} \mathcal{L} \times_{\mathcal{T}} \mathcal{C} &= (\underbrace{AA, AB, BB}, \underbrace{AC, BC}, \underbrace{AT, BT}) \times_{(\underbrace{TC,SC},\underbrace{AC,BC},\underbrace{AT,BT})} (\underbrace{TC, SC}, \underbrace{AC, BC}, CC) \\
(-1)^{(bb)(ac)} \mathcal{P} &= (\underbrace{AA, AB, BB}, \underbrace{AC, BC}, CC) \\
\mathcal{P} &= (AA, AB, AC, BB, BC, CC)
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{B} \times_{\mathcal{T}} \mathcal{C} &= (SB, SC, AB, AC, BB, BC, AT, BT) \times_{(TC,SC,AC,BC,AT,BT)} (TC, SC, AC, BC, CC) \\
(-1)^{(sc)(ab+bb)+(bb)(ac)} \mathcal{B} \times_{\mathcal{T}} \mathcal{C} &= (\underbrace{SB, AB, BB}, \underbrace{SC, AC, BC}, \underbrace{AT, BT}) \times_{(TC,\underbrace{SC,AC,BC},\underbrace{AT,BT})} (TC, \underbrace{SC, AC, BC}, CC) \\
(-1)^{(sc)(ab+bb)+(bb)(ac)} \mathcal{R} &= (\underbrace{SB, AB, BB}, \underbrace{SC, AC, BC}, CC) \\
\mathcal{R} &= (SB, SC, AB, AC, BB, BC, CC) \\
\mathcal{A} \times_{\mathcal{S}} \mathcal{R} &= (AA, \underbrace{AB, AC}, \underbrace{AT, AS}) \times_{(\underbrace{SB,SC},\underbrace{AB,AC},\underbrace{AT,AS})} (\underbrace{SB, SC}, \underbrace{AB, AC}, \underbrace{BB, BC, CC}) \\
\mathcal{P} &= (AA, \underbrace{AB, AC}, \underbrace{BB, BC, CC})
\end{aligned}
$$

Table 10.1: Confirming that the Binary Full Formula implies Mixed Associativity

left of BB. Since AC and BB appear together in an even number of products, however — to wit, in the two products $\mathcal{B}$ and $\mathcal{P}$ — reversing this choice would affect only the surface form of the Ternary Full Formula, not its underlying import.

To prove that the Ternary Full Formula 10-32 is correct with the products in the order that we have chosen, we apply the Binary Full Formula twice to evaluate the left-associated product $\big(\mathcal{A} \times_{\mathcal{S}} \mathcal{B}\big) \times_{\mathcal{T}} \mathcal{C}$. The resulting calculation is shown in the top half of Table 10.1. To save space, that table denotes direct products simply by listing their factor spaces; that is, we write (X, Y, Z) rather than $X \times Y \times Z$. We also abbreviate $\dim(AT)$ as (at), and similarly for the other subspaces. Note that we have to reorder some of the subspaces temporarily, in order to apply the Binary Full Formula. The horizontal braces indicate the grouping of subspaces into blocks for each application of the Binary Full Formula. And $\mathcal{L}$ denotes the result of the inner binary fiber product.

The bottom half of Table 10.1 verifies the Ternary Full Formula a second time, this time using the right-associated product $\mathcal{A} \times_{\mathcal{S}} \big(\mathcal{B} \times_{\mathcal{T}} \mathcal{C}\big)$. So $\mathcal{R}$ denotes the result of that inner binary fiber product. The two halves of Table 10.1 are not quite symmetric, because of our arbitrary choice to write AC to the left of BB in the Ternary Full Formula. But the two halves do get the same result, and this confirms that the Binary Full Formula (together with the Isomorphism Axiom) does indeed imply the Axiom of Mixed Associativity — the last of our other axioms.

Exercise 10-34 What is the *Binary Full Back-Formula*, the analog of the Binary Full Formula that would result if we replaced the Concatenate Axiom with the Concatenate-Backwards Axiom?

Answer: Swapping the left and right factor spaces in the Binary Full Formula gives

$$(SB \times AB \times BB) \times_{\overleftrightarrow{(SB \times AB \times AS)}} (AA \times AB \times AS) = AA \times AB \times BB.$$

But restoring alphabetic order among the names of the subspaces then requires that we swap the name AA with BB and the name AS with SB, leading to the final formula:

$$(AS \times AB \times AA) \times_{\overleftrightarrow{(AS \times AB \times SB)}} (BB \times AB \times SB) = BB \times AB \times AA.$$

10.4 Towards an n-ary Full Formula

In verifying the mixed associativity of the Binary Full Formula 10-23, we found it helpful to write out the Ternary Full Formula 10-32. Suppose that we wanted to write out an explicit n-ary Full Formula, for some larger n — a single formula that encodes the proper orientations of all n-ary transverse mixed fiber products. We would face two challenges.

First, to show that such a formula actually has the universality implied by the adjective "full", we need a result from linear algebra — essentially, a structure theorem for zigzags. We proved that theorem for binary zigzags by explicit linear algebra in Proposition 10-24. But extending that result to the n-ary case is not trivial. Indeed, we haven't yet extended it even to the ternary case. Fortunately, the theory of quiver representations comes to our rescue. We show, in Section 10.4.1, that the necessary structure theorem for n-ary zigzags is a corollary of Gabriel's Theorem, in the theory of quiver representations.

The other key ingredient that we need, in order to write out an explicit n-ary Full Formula, is an appropriate partial order. Each space that appears in that Full Formula will be written as the direct product of a sequence of subspaces, and we need to know an appropriate left-to-right ordering for the subspaces in that product. In Section 10.4.2, we analyze the constraints on such a partial order and we construct an explicit total order that satisfies all of the required constraints.

10.4.1 The indecomposable summands of a zigzag

In Proposition 10-24, we showed that every binary problem instance is isomorphic to an instance of the special type to which the Binary Full Formula applies. Recall that the proof had two steps. We first did explicit linear algebra to decompose the three linear spaces A, B, and S as direct sums of subspaces on which the factor maps f and g behave in very simple ways. We then adjusted the orientations of the resulting subspaces so that the orientations of the spaces A, B, and S came out properly. Adjusting the orientations turns out to be easy, even in the n-ary case. But the linear algebra required in the n-ary case is fairly subtle. Fortunately, the recent theory of *quiver representations* [4, 5, 14] has made dramatic progress on these subtle issues of linear algebra. The quiver Q_n that corresponds to an n-ary zigzag turns out be a quiver of *finite type*, so an appeal to Gabriel's Theorem will give us all that we need to analyze the structure of zigzags.

Define Q_n to be the directed graph that consists of a path of length $2n$ whose edges alternate between pointing forward and pointing backward:

$$\underset{D_1}{\circ} \longrightarrow \underset{C_1}{\circ} \longleftarrow \underset{D_2}{\circ} \longrightarrow \cdots \longrightarrow \underset{C_{n-1}}{\circ} \longleftarrow \underset{D_n}{\circ}$$

The names D_i and C_i are intended to suggest "domain" and "codomain", by the way. Since a quiver is just another name for a directed graph, we shall also refer to the graph Q_n as a quiver, the *n-ary alternating-path quiver*. An n-ary zigzag is, in other language, simply a representation of this quiver Q_n. In such a representation, we associate a linear space with each vertex, say the factor space A_i with the domain vertex D_i and the base space S_i with the codomain vertex C_i, for each i; and we associate a linear map with each edge, the linear map $f_i\colon A_i \to S_i$ with the i^{th} forward edge and the linear map $g_{i+1}\colon A_{i+1} \to S_i$ with the i^{th} backward edge, thus arriving at the n-ary zigzag

$$A_1 \xrightarrow{f_1} S_1 \xleftarrow{g_2} A_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} S_{n-1} \xleftarrow{g_n} A_n.$$

Given two n-ary zigzags Z_1 and Z_2, we construct their *direct sum* $Z = Z_1 \oplus Z_2$ by taking the direct sums of corresponding linear spaces and corresponding linear maps. The *all-zero zigzag*, the zigzag all of whose spaces are zero-dimensional and all of whose linear maps are identically zero, is an identity element of this direct-sum operator. A zigzag is *indecomposable* when it is not isomorphic to a direct sum of nonzero zigzags. The key to understanding the linear algebra of zigzags is to describe all of the isomorphism classes of indecomposable zigzags. Fortunately, the theory of quiver representations gives us that description.

The quiver Q_n is of finite type; that is, its Tits form [4, 14] is positive definite. So Gabriel's Theorem tells us that there are precisely $\binom{2n}{2}$ isomorphism classes of indecomposable representations of Q_n, one such isomorphism class for each of the $\binom{2n}{2}$ subpaths of Q_n. For example, consider the subpath of Q_n that goes from C_1 to D_3:

$$\underset{C_1}{\circ} \longleftarrow \underset{D_2}{\circ} \longrightarrow \underset{C_2}{\circ} \longleftarrow \underset{D_3}{\circ}$$

The indecomposable representations of Q_n that correspond to this subpath are all isomorphic, with the following structure: They associate 1-dimensional linear spaces with each of the vertices in the subpath, that is, with the four vertices C_1, D_2, C_2, and D_3; and they associate zero-dimensional spaces with the rest of the vertices of Q_n. There are then three linear maps for which both the domain and codomain are 1-dimensional: the forward map from D_2 to C_2 and the backward maps from D_2 to C_1 and from D_3 to C_2. For the representations in this equivalence class, those three linear maps are bijections. All of the other linear maps are identically zero, as they must be, including the forward map from $D_1 = \Diamond$ to C_1 and the forward map from D_3 to $C_3 = \Diamond$.

Theorem 10-35 (Structure of Zigzags) *Fix an arity $n \geq 1$. Every n-ary zigzag Z can be decomposed, in an essentially unique way, as the direct sum of a finite number of indecomposable n-ary zigzags, each of which corresponds to a subpath of the quiver Q_n, as discussed above. The structure of Z can thus be described, up to isomorphism, by specifying how many of these indecomposable summands lie in each of the $\binom{2n}{2}$ possible isomorphism classes. Those classes can be grouped into four types, according as the corresponding subpath runs*

Type DD**:** *from D_i to D_j, for some $i \leq j$;*

Type DC**:** *from D_i to C_j, for some $i \leq j$;*

Type CD**:** *from C_i to D_j, for some $i < j$; or*

Type CC**:** *from C_i to C_j, for some $i \leq j$.*

Proof This is Gabriel's Theorem [4, 14] applied to the particular quiver Q_n, which is of finite type. The $\binom{2n}{2}$ isomorphism classes correspond to the $\binom{2n}{2}$ nonnegative dimension vectors at which the Tits form takes on the value $+1$. □

Theorem 10-35 is a structure theorem for zigzags, and we shall use it to verify that an n-ary Full Formula is actually "full". Before we do that, however, we need to consider the role that transversality plays in Theorem 10-35.

Proposition 10-36 *An indecomposable n-ary zigzag is transverse just when it is not of Type* CC.

Proof Recall that an n-ary zigzag $A_1 \xrightarrow{f_1} S_1 \xleftarrow{g_2} A_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{n-1}} S_{n-1} \xleftarrow{g_n} A_n$ is transverse just when its difference map $\Delta\colon A_1 \times \cdots \times A_n \to S_1 \times \cdots \times S_{n-1}$ is surjective, where that difference map is given by

$$\Delta(\alpha_1, \ldots, \alpha_n) := \bigl(g_2(\alpha_2) - f_1(\alpha_1),\ \ldots,\ g_n(\alpha_n) - f_{n-1}(\alpha_{n-1})\bigr).$$

If a zigzag is indecomposable of Type CC, the number of base spaces S_i that are 1-dimensional exceeds by one the number of factor spaces A_i that are 1-dimensional; so the difference map Δ can't possibly be surjective.

If a zigzag is indecomposable of Type DC, its difference map is bijective, as we can check by working from right to left. For example, if its 1-dimensional spaces run from the vertex D_i to the vertex C_j, for some $i \le j$, then any value for σ_j determines a unique corresponding value for α_j, after which that value, together with any value for σ_{j-1}, determines a unique value for α_{j-1}, and so forth. The difference maps for indecomposable zigzags of Type CD are similarly bijective, as we can check by working from left to right.

Finally, if a zigzag is indecomposable of Type DD, its difference map fails to be injective, but is surjective — as we can check by working in either direction. □

Proposition 10-37 *If an n-ary zigzag Z is a direct sum, say $Z = \bigoplus_i Z_i$, then Z is transverse just when all of the summands Z_i are transverse.*

Proof The difference map of Z is the direct product of the difference maps of the Z_i, and a direct product of maps is surjective just when each factor is surjective. □

Corollary 10-38 *An n-ary zigzag is transverse just when all of its indecomposable summands are of types other than Type* CC.

We can now verify the "fullness" of an n-ary Full Formula. We shall state the result only for the ternary case, $n = 3$; but the same ideas suffice for all n.

Proposition 10-39 *Every ternary problem instance $A \xrightarrow{f} S \xleftarrow{g} B \xrightarrow{h} T \xleftarrow{k} C$ is isomorphic to an instance of the special form to which the Ternary Full Formula applies, that special form being as described in Formula 10-32.*

Proof As in our analysis of the binary case in Proposition 10-24, the proof has two parts. The hard part is the linear algebra that breaks up all of the factor spaces and all of the base spaces into appropriate subspaces. It is then straightforward to adjust the orientations of those subspaces as needed.

We deal with the linear algebra by applying Theorem 10-35 to the given ternary problem instance, thereby decomposing it as the direct sum of indecomposable zigzags of the fifteen classes shown in Table 10-31. Of those fifteen classes, the three classes SS, ST, and TT cannot arise, since a ternary problem instance is assumed to be transverse. We group the indecomposable summands of the twelve remaining classes by their class. The direct sum of the summands of class SC, for example, then gives us the subspaces S_{SC}, B_{SC}, T_{SC}, and C_{SC}, and similarly for the other eleven classes.

It remains only to deal with the orientations. We begin by making an arbitrary choice of orientation for each of the twelve classes. For example, for the class SC, we choose some orientation for the subspace S_{SC}, and we carry that orientation forward to orient B_{SC}, T_{SC}, and C_{SC}. We then reverse some of our arbitrary choices, if necessary, in order to arrange that the three factor spaces A, B, and C and the two base spaces S and T all emerge with the proper orientations. We fix up the base spaces first. If the orientation of S is currently wrong, we reverse the subspaces A_{AS} and S_{AS}. Similarly, if T is currently wrong, we reverse B_{BT} and T_{BT}. We can then fix up each factor space independently, with no interference. For example, if B is currently wrong, we simply reverse B_{BB}. □

10.4.2 The poset underlying a Full Formula

Our quiver-based analysis of the structure of a zigzag has demonstrated that every transverse n-ary zigzag can be decomposed in such a way that an n-ary Full Formula will apply. To write out an explicit n-ary Full Formula, however, we need one further ingredient: We need to know how to order the various subspaces in each of its direct products. In the ternary case, for example, we ordered the six direct products in Formula 10-32 as specified by the chains in Poset 10-33.

In the n-ary case, we need a poset whose elements are the isomorphism classes of indecomposable representations of an n-ary zigzag. Of those $\binom{2n}{2}$ classes, however, the $\binom{n}{2}$ classes that are of Type CC will not arise, since our zigzags are transverse. So our poset will have $\binom{2n}{2} - \binom{n}{2}$ elements.

What are our goals for this poset? To be useful, it must impose a linear order on all of the isomorphism classes that appear in any single product of our n-ary Full Formula. But we also want our Full Formula to be correct, that is, to impose the Proper Orientation on every fiber product. We can guarantee that correctness by imposing simple combinatorial conditions on the poset.

Definition 10-40 Fixing $n \geq 1$, consider a poset whose elements are the $\binom{2n}{2} - \binom{n}{2}$ isomorphism classes of indecomposable, transverse representations of the quiver Q_n. We call such a poset *admissible* when it meets the following six conditions:

1. For each vertex in the quiver Q_n, the poset imposes a linear order on those classes that include that vertex.

2. The poset also imposes a linear order on all of the classes of Type DD. Note that those classes are the ones that appear, on the right-hand side of the Full Formula, in the expression for the fiber product.

3. If the leftmost vertex in a class α_{DD} of Type DD also belongs to a class β_{CD} of Type CD, then the partial order must have $\beta_{\mathsf{CD}} \prec \alpha_{\mathsf{DD}}$.

4. Symmetrically, if the rightmost vertex in a class α_{DD} of Type DD belongs to a class β_{DC} of Type DC, then we must have $\alpha_{\mathsf{DD}} \prec \beta_{\mathsf{DC}}$.

5. If a class α_{CD} of Type CD and a class β_{DC} of Type DC intersect, but are not nested one inside the other, then we must have $\alpha_{\mathsf{CD}} \prec \beta_{\mathsf{DC}}$. An equivalent way to describe this situation is to say that the two classes α_{CD} and β_{DC} share an odd number of vertices.

6. Given two classes α_{DD} and β_{DD}, both of Type DD, if every vertex of α_{DD} lies strictly to the left of every vertex of β_{DD}, then we must have $\alpha_{\mathsf{DD}} \prec \beta_{\mathsf{DD}}$.

In the binary case $A \times_S B$, each of the last four conditions boils down to a single ordering constraint. Condition 3 says $SB \prec BB$, Condition 4 says $AA \prec AS$, Condition 5 says $SB \prec AS$, and Condition 6 says $AA \prec BB$. Note that these four constraints are precisely the four that we saw to be necessary in Section 10.3.3.

We shall show, in a moment, that the six conditions in Definition 10-40 suffice to ensure that the resulting n-ary Full Formula is correct. But let's first make the easy observation that these six conditions can be simultaneously satisfied.

Proposition 10-41 *Admissible partial orders exist, for every $n \geq 1$.*

Proof We satisfy Conditions 1 and 2 by choosing our partial order to be total. In Conditions 3 through 5, note that a class of Type CD is always required to be smaller than some class of some other type, while a class of Type DC is always required to be larger than some class of some other type. We can hence satisfy those conditions as follows: We put all of the classes of Type CD first in our order, ordered arbitrarily among themselves. And we put all of the classes of Type DC last, again ordered arbitrarily among themselves. It remains only to insert the classes of Type DD into the middle, between the CD's and the DC's, in some order that satisfies Condition 6. For example, we could describe each class of Type DD by the pair (leftmost vertex, rightmost vertex) and then sort those pairs lexicographically. $\square$

Proposition 10-42 *The n-ary Full Formula that is constructed from any admissible partial order is correct, in the sense that it determines the Proper Orientation for every transverse, n-ary fiber product.*

Proof To show that the resulting n-ary Full Formula is correct, we use our axioms to reduce its left-hand side to coincide with its right-hand side.

Step 1: Canceling the classes of Types CD and DC We start by tackling the isomorphism classes of Type CD and DC, working inductively from the shortest such classes toward the longer ones. We shall eliminate each such class in turn from our Full Formula — in particular, from that formula's left-hand side.

Choose one of the shortest remaining classes. Let's say that it is the class σ_{CD} of Type CD; if it is of Type DC, we treat it symmetrically. The class σ_{CD} includes an even number of vertices, and we view each codomain vertex as paired with the following domain vertex. Let c and d denote one such adjacent pair of vertices, included within σ_{CD}; we shall treat each such pair separately. We are going to use the Right Partial Identity Formula 10-11 to remove the factor σ_{CD} from the left-hand-side products associated with the vertices c and d.

We first exploit mixed associativity to insert parentheses into our Full Formula in such a way that one of the innermost fiber products to be computed is the binary product that has vertex c as its base space and vertex d as its right-hand factor space. Note that the direct-product expressions associated with each of these vertices will include a factor corresponding to σ_{CD}. If those two factors were both leftmost in their products, we could apply the Right Partial Identity Formula immediately to remove those two factors.

Typically, of course, there will be factors to the left of σ_{CD} in the product for the vertex c, and also to the left of σ_{CD} in the product for the vertex d. Some of those leftward factors may be shared between the two products. In fact, we claim that all such leftward factors must be shared.

Consider some factor, say τ, that lies to the left of σ_{CD} in the product for the vertex d. If the factor τ occurs at all in the product for c, it must lie to the left of σ_{CD} there as well, since the same partial order linearly ordered both products. Can it be that τ does not occur at all in the product for c? If so, the class τ must have d as its leftmost vertex; so τ must be either of Type DD or of Type DC.

Suppose first that τ_{DD} is of Type DD. By Condition 3, since d is the leftmost vertex in τ_{DD} and d lies also in σ_{CD}, the partial order must have $\sigma_{\mathsf{CD}} \prec \tau_{\mathsf{DD}}$. But that contradicts the leftward location of τ_{DD}.

On the other hand, suppose that τ_{DC} is of Type DC. By induction, we have already removed from our Full Formula all factors corresponding to classes that are shorter than σ_{CD}. So the class τ_{DC} is at least as long as σ_{CD}. Since τ_{DC} starts to the right of where σ_{CD} starts, the classes σ_{CD} and τ_{DC} cannot be nested. Condition 5 therefore applies, telling us that $\sigma_{\mathsf{CD}} \prec \tau_{\mathsf{DC}}$, which is again a contradiction.

So every factor that occurs to the left of σ_{CD} in the product for the vertex d occurs also in the product for c. What about the converse?

Suppose that ρ is a factor that lies to the left of σ_{CD} in the product for c; could it be that ρ does not occur in the product for d? The rightmost vertex of ρ would then have to be c; so ρ would have to be of Type DC. (There is no second choice because Type CC doesn't occur.) Since ρ_{DC} is at least as long as σ_{CD}, those two can't be nested. By Condition 5, we conclude that $\sigma_{\mathsf{CD}} \prec \rho_{\mathsf{DC}}$, which is again a contradiction.

We have now shown that the factors that lie to the left of σ_{CD} in the product for the vertex c are identical to those that lie to the left of σ_{CD} in the product for d. We can swap σ_{CD} to the left over each of these factors in turn. Since each such swap happens twice, once in the formula for c and once in the formula for d, any factor of -1 that might be introduced cancels out. Having done these swaps, the factor σ_{CD} is now leftmost in the products for both c and d, so we can use the Right Partial Identity Formula to eliminate it. Note that this swapping puts our formula, temporarily, into a state where some of its products are not ordered according to the controlling partial order. But, once we eliminate the two instances of σ_{CD} with the Right Partial Identity Formula, that problem goes away.

Step 2: Converting to lexicographic order When the induction in Step 1 has been completed, all of the factors that remain in our Full Formula, on the left-hand side as well as on the right, correspond to classes of Type DD. By Condition 2, our partial order restricts to total order on all of the classes of Type DD. Our next goal is to replace that total order, one swap at a time, with lexicographic order.

A pair of classes of Type DD is currently *inverted* if the ordering relationship between those classes in the current total order is the opposite of their lexicographic relationship. Our strategy will be to eliminate each such inverted pair by performing a single swap. We don't want that swap to affect any other ordering relationships, however; so we restrict ourselves to swapping inverted pairs that are adjacent in the current total order. If there are no inverted pairs that are currently adjacent, then every adjacent pair is in lexicographic order. By transitivity, the current order must be entirely lexicographic, so we are done with Step 2.

If not, choose some inverted pair that is currently adjacent: Say that τ_{DD} and σ_{DD} are two classes of Type DD that are adjacent in the current total order, that $\tau_{\mathsf{DD}} \prec \sigma_{\mathsf{DD}}$ in that total order, but that $\sigma_{\mathsf{DD}} \prec \tau_{\mathsf{DD}}$ in lexicographic order. Our plan is to swap σ_{DD} to appear before τ_{DD}.

The only way in which we are changing any ordering relationships is to swap inverted pairs into lexicographic order. So any pair that is currently inverted was also originally inverted. But the original total order on the classes of Type DD was required to satisfy Condition 6. Since τ_{DD} preceded σ_{DD} in that original order, it cannot have been the case that σ_{DD} preceded τ_{DD}; so, by the contrapositive of Condition 6, it cannot be the case that every vertex of σ_{DD} lies to the left of every vertex of τ_{DD}. But we also know that σ_{DD} precedes τ_{DD} in lexicographic order. Because of the way that lexicographic order works, it cannot be the case that every vertex of σ_{DD} lies to the right of every vertex of τ_{DD}. The only remaining option is that the vertex sets of σ_{DD} and τ_{DD} intersect.

Two classes of Type DD that intersect can intersect in various ways: One may be nested inside the other, or they may overlap like shingles. In any case, however, the size of their intersection is always odd. So suppose that we rewrite the left-hand and right-hand sides of our Full Formula to adjust for swapping σ_{DD}

to precede τ_{DD}. We need to swap adjacent factors corresponding to σ_{DD} and τ_{DD} in an odd number of products on the left-hand side, one for each vertex in their intersection. But we also need to swap them in the single product on the right-hand side. So the total number of swaps is even, and we can alter our Full Formula to account for this swap without introducing a factor of -1.

Step 3: Canceling the superfluous factors of Type DD As a result of Step 2, the total order on the classes of Type DD the controls our current Full Formula is now lexicographic. And that makes it easy to cancel the superfluous factors that remain. To see the pattern, here is what the left-hand side of a quaternary fiber product would be at the start of these final simplifications, given that the controlling order is lexicographic:

$$\begin{aligned}(AA, AB, AC, AD) \times_{(AB,AC,AD)} &(AB, AC, AD, BB, BC, BD)\\ \times_{(AC,AD,BC,BD)} &(AC, AD, BC, BD, CC, CD)\\ \times_{(AD,BD,CD)} &(AD, BD, CD, DD).\end{aligned}$$

(We have separated the factors in each direct product here with commas, rather than with $\times$'s, just to save space.) Consider each base space. We can use the Right Partial Identity Formula to cancel each of the factors of that base space, from left to right, with the corresponding factors at the start of the following factor space. All of the base spaces in what remains are simply $\Diamond$, so we can convert the fiber products that remain into direct products using the Concatenate Axiom. And what results from that is precisely the direct product of all classes of Type DD, in lexicographic order — which agrees with the right-hand side:

$$(AA, AB, AC, AD, BB, BC, BD, CC, CD, DD)$$

So the Full Formula that we started with was indeed correct. □

Exercise 10-43 If a partial order is going to guide us in writing out an n-ary Full Formula, that order certainly has to satisfy Conditions 1 and 2 in Definition 10-40. Proposition 10-42 shows that Conditions 3 through 6 then suffice to guarantee that the resulting n-ary Full Formula will be correct. In the other direction, show that Conditions 3 through 6 are also necessary. In particular, given a partial order that violates even just one of the ordering constraints in Conditions 3 through 6, find a fiber product that the resulting n-ary Full Formula will orient improperly.

Hint: The arguments are similar to those in Exercises 10-29 and 10-30. In each case, we can choose the subspaces corresponding to the two misordered classes to be 1-dimensional, while choosing all other subspaces to be $\Diamond$.

Bibliography

[1] Julien Basch, Leonidas J. Guibas, G. D. Ramkumar, and Lyle Ramshaw. Polyhedral tracings and their convolutions. *Algorithms for Robotic Motion and Manipulation: 1996 Workshop on the Algorithmic Foundations of Robotics*, edited by Jean-Paul Laumond and Mark Overmars, A. K. Peters (1997), pages 171–184.

[2] Edward Bierstone and Pierre D. Milman. Semianalytic and subanalytic sets. *Publications Mathématiques* of the Institut des Hautes Études Scientifiques **67** (1988), pages 5–42.

[3] Garrett Birkhoff. *Lattice Theory,* third edition. American Mathematical Society (1967), pages 65 and 66, particularly Exercise 8 on page 66.

[4] Harem Derksen and Jerzy Weyman. Quiver representations. *Notices of the AMS* **52** #2 (2005), pages 200–206.

[5] Peter Gabriel and Andrei V. Roiter. *Representations of Finite-Dimensional Algebras,* second printing. Springer-Verlag (1997), particularly Chapter 7.

[6] Leo Guibas, Lyle Ramshaw, and Jorge Stolfi. A kinetic framework for computational geometry. *IEEE 24th Annual Symposium on the Foundations of Computer Science* (1983), pages 100–111.

[7] John Philip Huneke. Mountain climbing. *Transactions of the American Mathematical Society* **139** (1969), pages 383–391.

[8] Serge Lang. *Differential and Riemannian Manifolds*. Springer-Verlag (1995).

[9] Page 21.

[10] Pages 29 and 30.

[11] Saunders Mac Lane. *Categories for the Working Mathematician,* second edition. Springer-Verlag (1998).

[12] Page 196.

[13] G. D. Ramkumar. *Tracings and their Convolution: Theory and Applications*. PhD thesis, Stanford University (1998).

[14] Alistair Savage. Finite-dimensional algebras and quivers. *Encyclopedia of Mathematical Physics*, volume 2, edited by J. P. Francoise, G. L. Naber, and S. T. Tsou, Oxford:Elsevier (2006), pages 313–320. Also available as “arXiv:math/0505082v1”.

[15] Pierre Schapira. Operations on constructible functions. *Journal of Pure and Applied Algebra* **72** (1991), pages 83–93.

[16] James V. Whittaker. A mountain-climbing problem. *Canadian Journal of Mathematics* **18** (1966), pages 873–882.

Index