Chapter 2
Parametrization

This chapter develops several models of topological spaces associated to collections of measurements that

1. Parametrize the context in which the measurements are taken,
2. Set the stage for the understanding when measurements are consistent, and
3. Are convenient to implement in applications.

Local signals are collections of measurements which are related by their spatial, temporal, or contextual proximity. This leads one to study the sets in which "proximity" has a consistent mathematical meaning, namely *topological spaces*. The general definition of a topological space, while elegant and intellectually efficient, sadly admits many pathologies. (Indeed, it admits so many pathologies that at least one book, Steen and Seebach (1978), has been written about them!) Because of this, we will study several restricted classes of topological spaces which admit fewer pathologies. Of course, there could be a situation of interest which is not admitted in these classes, but the author has not yet found an example of such a situation. Additionally, these classes of spaces are well-suited to algorithmic manipulation, and therefore play an important role throughout this book.

The primary class of topological spaces that are used in applications are *cell complexes*. These include some *CW complexes* and all *simplicial complexes* as shown in Fig. 2.1. From a computational perspective, simplicial complexes are quite convenient to manipulate, though in high-dimensional settings they can incur a large memory footprint. It is for this reason that others, for instance Kaczynski et al. (2004), also consider cubical complexes. However, most of the cases of interest for signal processing use lower dimensions, so we will avoid the additional technical overhead of discussing cubical complexes.

When more structure is available in the problem, it is sometimes useful to instead study *manifolds*, which are spaces that are locally like Euclidean space. Conveniently, every manifold can be given the structure of a cell complex, and every compact manifold can be given the structure of a CW complex. Unlike cell complexes and CW complexes, manifolds admit a general notion of differential calculus. Because manifolds admit both a differential and a combinatorial interpretation, manifolds

M. Robinson, *Topological Signal Processing*,
Mathematical Engineering, DOI: 10.1007/978-3-642-36104-3_2,

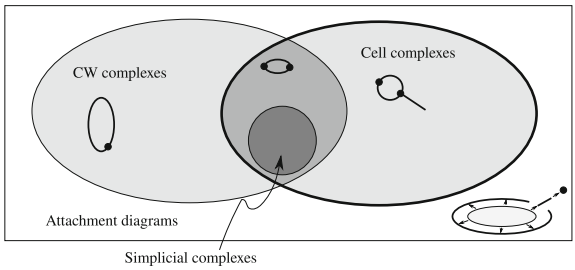

Fig. 2.1 Some kinds of topological spaces that are important in applications

can be useful in cementing the relationship between a physical signal model (arising from differential equations) and the algorithmic, combinatorial models we develop in this book.

2.1 Abstract Spaces

We begin our study of restricted topological spaces with *CW complexes*. These spaces are assembled from disks glued along their boundaries in a particular way. In the end, we will focus on related spaces, called *cell complexes*, because these are the most natural place to study sheaves and local signals.

2.1.1 CW Complexes

In order to connect computations concerning the structure of a space to data, it is essential to have a constructive, combinatorial description of a space that arises from attaching simpler topological spaces together.

Definition 2.1 Suppose that X and Y are topological spaces and that $f : A \subseteq X \to Y$ is a continuous function from a subspace of X. Let $Z = (X \backslash A) \sqcup Y$, and define the function $g : X \sqcup Y \to Z$ by

$$g(p) = \begin{cases} p & \text{if } p \in Y \text{ or } p \notin A \\ f(p) & \text{if } p \in A \end{cases}$$

The *adjunction space* is the topological space $(Z, \mathscr{U})$, whose topology $\mathscr{U}$ consists of all subsets of Z whose preimages are open in $X \sqcup Y$. We say that Z consists of X and Y *attached* along $A \sim f(A)$. (See Fig. 2.2 for a schematic of this construction.)

General attachments can be rather complicated, so we will instead consider spaces that arise from attaching disks to one another along their boundaries.

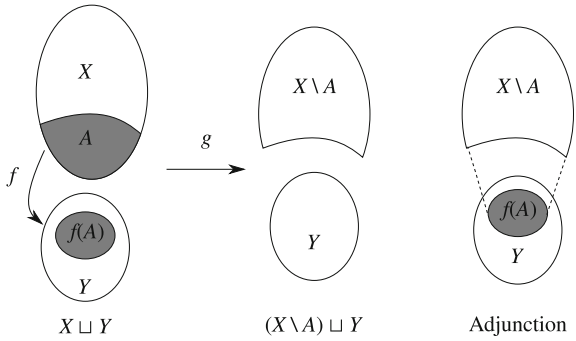

Fig. 2.2 An adjunction of two spaces

Definition 2.2 The n-dimensional *closed disk* is the closed subset $D^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 \leq 1\}$. The boundary of such a disk is the $n-1$-dimensional *sphere* $\partial D^n = S^{n-1}$, the closed subset $\{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\}$. (We will use the notation ∂A to represent the boundary of a set A.) Similarly, the n-dimensional *open disk* is the subset $\{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 < 1\}$.

Definition 2.3 Suppose that the topological space X^k consists of the (not necessarily disjoint) union of a collection $\mathscr{K}$ of disks of dimension at most k, and that ∂D^{k+1} is decomposed as a finite union $\partial D^{k+1} = A_1 \cup A_2 \cdots \cup A_n$. An *attaching map*[1] or *attachment* is a continuous function $\phi : \partial D^{k+1} \to X^k$, for which the image $\phi(A_i)$ of each A_i is a disk, and each such image is in the collection $\mathscr{K}$.

We then say that D^{k+1} is *attached* to X^k and to each of the disks $\phi(A_i)$, and will write $\phi(A_i) \rightsquigarrow D^{k+1}$. Observe that the symbol $\rightsquigarrow$ goes in the opposite direction as the attaching map itself, a convention that is useful when we study sheaves.

The concept of a *CW complex* captures the idea of inductively constructing a space from lower dimensions to higher dimensions using attaching maps, as suggested in Fig. 2.3.

Definition 2.4 Suppose that X^0 is a disjoint collection of points and that X^k is the adjunction space formed by attaching finitely many disks to X^{k-1} for $k = 1, \ldots, n$ as above. Then $X = X^n$ is a *finite CW complex* of dimension n, or simply a *CW complex*. The space X^k is called the *k-skeleton* of X.

Attachment constructions like these are powerful tools for understanding how a space is assembled from its parts. They play a similar role to exploded view diagrams (see Fig. 2.4) in assemblies. Because of this graphical connection, attachment constructions can be visualized with an *attachment diagram*, showing how the cells are attached, such as in Fig. 2.5. Each cell is shown and the links represent attachments (drawn from low dimensional cells to higher dimensional cells). Usually, we suppress any attachments that are compositions of other attachments.

[1] The term *map* will be used as shorthand for "continuous function" throughout the book.

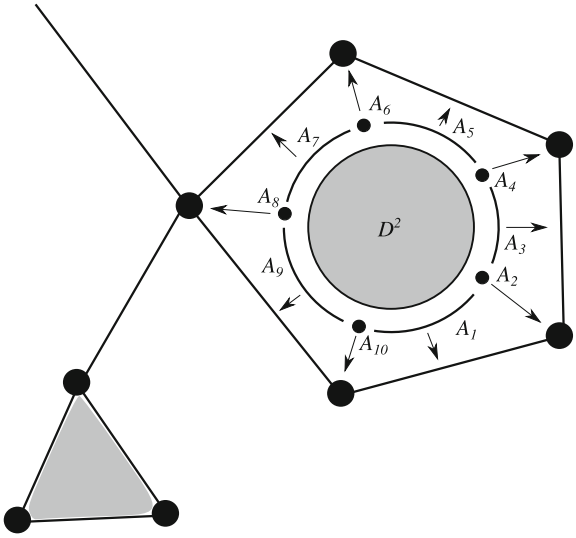

Fig. 2.3 Example of an attaching map (marked with *arrows*) of a disk D^2

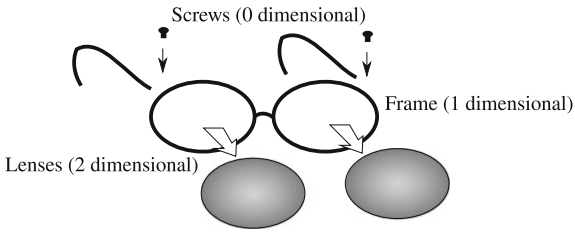

Fig. 2.4 An attachment construction is like an exploded view for a topological space

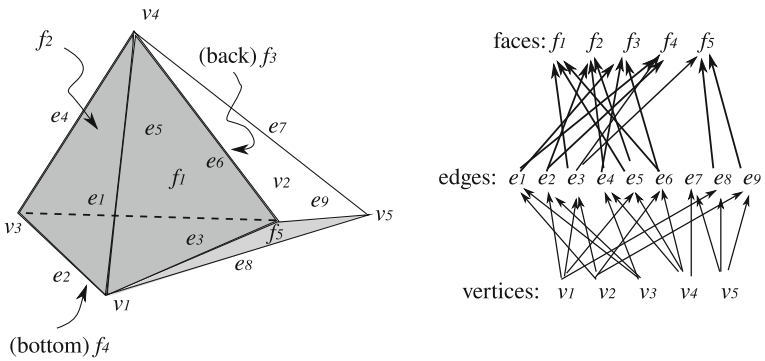

Fig. 2.5 An attachment construction (*left*) and its attachment diagram (*right*)

Remark 2.1 The attachment diagram of an attachment construction is a set that is partially ordered by the sequence of attachments. When we exhibit an attachment diagrams, we display its Hasse diagram.

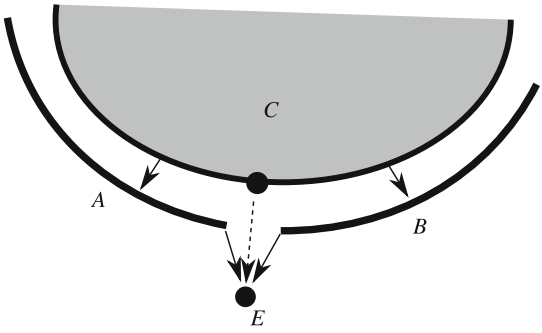

Fig. 2.6 The compatibility condition for attaching maps in CW complexes; the attaching map from the boundary of C to A also restricts to the attachment of A to E

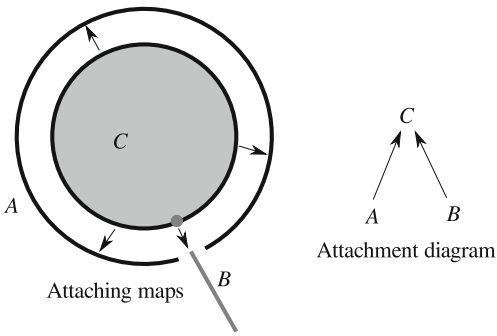

Fig. 2.7 An attachment construction that is not a CW complex. Note that B is an edge attached to a 2-dimensional disk C without a vertex between them

Because we assume that attaching maps are continuous, no new attachments between cells in X^k arise when we attach higher dimensional cells. As shown in Fig. 2.6, if two cells A and B are both attached to a higher dimensional cell C at the same place on C, then they must also be attached to the same *lower* dimensional cell E there also. If we remove the higher dimensional cell C, the attachment diagram retains the common attachments of A and B to E.

Remark 2.2 If the attaching maps are not assumed to be continuous, the resulting construction would not necessarily be an adjunction space. For example, the disks A and B in Fig. 2.7 would become detached from one another if the higher-dimensional cell C were to be removed. With the cell C present, the 1-skeleton of the attachment construction is a different topological space than with C removed.

Corollary 2.1 *The k-skeleton of a CW complex is also a CW complex.*

The boundaries of closed disks are CW complexes, and they can be easily assembled inductively. To start the induction, we begin by constructing a circle as a CW complex. This structure consists of two 1-cells (closed intervals) $e_1 = [0, 1]$ and $e_2 = [0, 1]$, attached to two 0-cells (points) v_1 and v_2 as shown in Fig. 2.8. Specifically, there are two attaching maps, namely

- $a_1 : \partial e_1 = \{0, 1\} \to X^0 = \{v_1, v_2\}$ given by $a_1(0) = v_1$ and $a_1(1) = v_2$, and

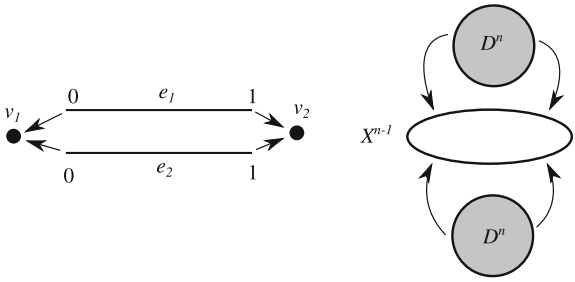

Fig. 2.8 CW complex structures for a circle (*left*) and sphere (*right*). (From a geometric perspective, this construction of S^2 looks more like a "whoopie cushion" than a sphere.)

- $a_2 : \partial e_2 = \{0, 1\} \to X^0 = \{v_1, v_2\}$ given by $a_2(0) = v_1$ and $a_2(1) = v_2$.

To construct an n-dimensional sphere, we assume that the $n-1$-skeleton X^{n-1} consists of a single $n-1$ dimensional sphere as a CW complex. We then attach two n-dimensional disks by attaching the boundary of each to X^{n-1} with a homeomorphism as shown in Fig. 2.8.

Other constructions are possible as well. For instance, the n-dimensional *torus* is constructed by attaching opposite faces of the n-dimensional cube to each other. This is not quite as constrained of a definition as it might first seem; how precisely the attachment is made does matter. Consider the two-dimensional situation, in which opposite sides of a square are attached to each other. If, for instance, $(0, x)$ is attached to $(1, x)$ and $(x, 0)$ is attached to $(x, 1)$, the resulting space is a torus. However, if instead we consider the CW complex that arises from attaching $(0, x)$ to $(1, x)$ and $(x, 0)$ to $(1 - x, 1)$, the resulting space is called a *Klein bottle* and is quite different. The Klein bottle is a space which cannot be homeomorphically mapped into a 2-dimensional subset of $\mathbb{R}^3$. A standard projection of its embedding[2] into $\mathbb{R}^4$ to $\mathbb{R}^3$ is shown in Fig. 2.9.

Exercise 2.1 The *real projective n-plane* $\mathbb{RP}^n$ arises as a quotient of the closed n-disk D^n by the equivalence $v \sim -v$ for each $v \in \partial D^n$. Construct a CW complex structure for $\mathbb{RP}^n$.

While flexible, CW complexes have the disadvantage that they are sometimes too constrained. This is especially the case when studying graphs; a CW complex representation of a graph requires both endpoints of every edge to be connected to a vertex. Sometimes, however, it is more descriptive to remove a vertex from the end of an edge. This might signal that the edge's connection on that end is not useful to consider, or might indicate a connection outside of the region being considered. The appropriate generalization is that of a *cell complex*. A cell complex is a either (1) a CW complex whose attaching maps are embeddings or (2) a space that becomes a CW complex after being compactified, and whose attaching maps are embeddings even once the space has been compactified.

[2] An *embedding* is an injective continuous map $f : X \to Y$ whose image $f(X)$ is homeomorphic to its domain X.

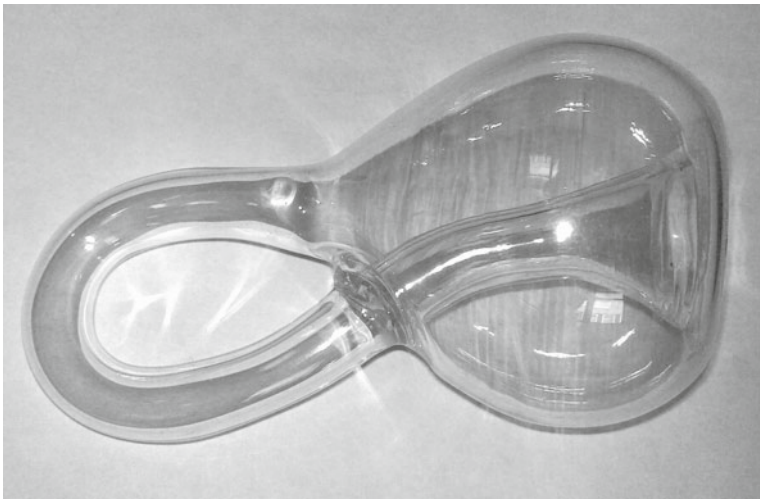

Fig. 2.9 Projection of an embedded Klein bottle into $\mathbb{R}^3$

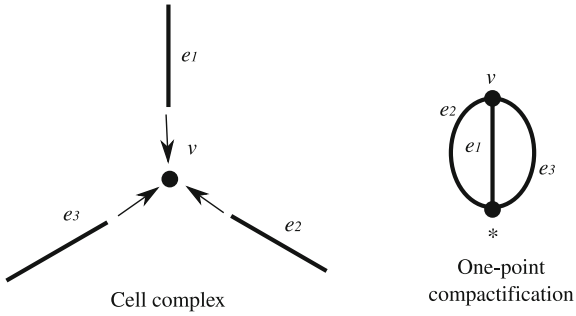

Fig. 2.10 A cell complex (*left*) that is not a CW complex. Its one-point compactification (*right*) has the structure of a CW complex

Definition 2.5 A *cell complex* consists of a topological space X and a collection of disjoint *open* disks $\{c_k\}$, each of which is a subspace of X, such that if we add an additional *external vertex* $*$ and form $\{\overline{c_k}\} \cup \{*\}$, these are the cells of a CW complex structure for the one-point compactification $X \cup \{*\}$ (see Proposition A.7 in the Appendix) of X, whose attaching maps are all embeddings. We call the open disks $\{c_k\}$ *open cells* whenever there is the possibility of confusion.

Figure 2.10 shows a cell complex that is not a CW complex, since only one endpoint of each edge is attached to a vertex. However, its one-point compactification is a CW complex.

The simpler space, consisting of a single open interval is not a cell complex, however. Its one-point compactification is a circle, but its only attaching map is not an embedding since both endpoints are attached to the same (new) point. This problem can be rectified by subdividing the open interval into two open intervals attached to a common point in the middle.

2.1.2 Cellular Maps and Homotopy

Cell complexes can be related to one another in much the same way that continuous maps relate topological spaces.

Definition 2.6 A continuous map $f : X \to Y$ between two cell complexes is called a *cellular map* if

1. For each cell $c \in X$, $f(c)$ is a cell in Y,
2. If $a \rightsquigarrow b$ is an attachment of two cells in X then $f(a) \rightsquigarrow f(b)$ is an attachment in Y, and
3. Given a cell $c \in X$ and two other cells $d, e \in f(c)$, then $f^{-1}(d) \cap \overline{c}$ is compact if any only if $f^{-1}(e) \cap \overline{c}$ is also compact.

It is worth noting that condition (3) is automatically satisfied if X and Y are CW complexes, and the last two conditions are automatically satisfied if X and Y are both abstract simplicial complexes (Sect. 2.2.1).

Topology is popularly understood to be the branch of mathematics that deals with deformations, particularly of images like the sequence shown in Fig. 2.11. This is an example of a *homotopy* from one image to another.

Definition 2.7 A *homotopy* between two maps $f : X \to Y$ and $g : X \to Y$ is a continuous function $H : X \times [0, 1] \to Y$ in which $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$ for all $x \in X$. We call f and g *homotopic* (written $f \simeq g$) if there is a homotopy between f and g.

The fact that a single topological space may be described by several different cell complex structures raises the question of how to tell if two cell complexes are the "same". One can ask whether two spaces are homeomorphic (or stronger, homeomorphic through a cellular map), but this is too stringent of a criterion. CW complexes are often best studied under a different criterion, called *homotopy equivalence*.

Definition 2.8 A *homotopy equivalence* between two topological spaces X and Y is a pair of maps $f : X \to Y$ and $g : Y \to X$ such that both $f \circ g$ and $g \circ f$ are homotopic to identity maps. We say that X and Y are *homotopy equivalent* if there exists a homotopy equivalence between them.

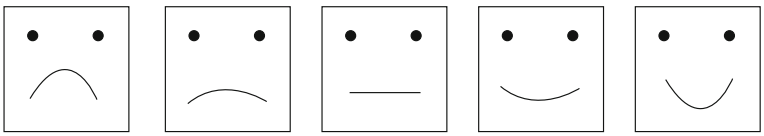

Fig. 2.11 A continuous deformation of one image into another

Exercise 2.2 Show that [0, 1] is homotopy equivalent to itself, by showing that the compositions of $f, g : [0, 1] \to [0, 1]$ given by

$$f(x) = \sqrt{3x} \text{ and } g(y) = 4y^2$$

are homotopic to the identity map $\mathrm{id}_{[0,1]}(x) = x$.

Rather than directly checking if the resulting topological spaces associated to an attachment construction are homotopic, it is usually easier to look for algebraic "homotopy invariants" that allow different spaces to be discriminated. If two spaces are homotopy equivalent, they have the same sets of invariants; so if the invariants associated to two spaces are different, then the spaces cannot be homotopy equivalent. However, the converse is not true; if two spaces have the same set of invariants, they may still not be homotopy equivalent.

For CW complexes, the invariants which discriminate between the most spaces are found in homotopy groups. Unfortunately, computing homotopy groups is computationally and theoretically very difficult, and there remain many open questions about how to effectively perform the computation at all. On the other hand, much of the complexity can be avoided by using weaker invariants called *homology* and *cohomology*. While easier to compute, homological invariants cannot discriminate between certain homotopy *in*equivalent spaces. Nevertheless, we will develop and use homology theory throughout the rest of the book because of its computational utility.

2.2 Representation of Spaces

As convenient as CW complexes are for constructing and manipulating spaces, they are remarkably difficult to represent completely in software. Because of this, it is useful to have more restricted descriptions of spaces that have the same notion of topology as CW complexes, but are more computationally convenient. From a theoretical point of view, one can switch between equivalent representations as appropriate.

2.2.1 Abstract Simplicial Complexes

The prototypical combinatorial model of a space is an *abstract simplicial complex*. It is extremely easy to store and manipulate an abstract simplicial complex in a computer because it can be represented using ordered lists.

Definition 2.9 An *abstract simplicial complex* X on a set A is a collection of ordered[3] lists of A that is closed under the operation of taking sublists. We call each element of X, which is an ordered collection of elements of A, a *simplex*. A simplex with $k+1$ elements is called a k-simplex or a k-face, and we call a 0-face a *vertex* and a 1-face an *edge*. A *simplicial map* $f : X \to Y$ between an abstract simplicial complex on A to an abstract simplicial complex on B is a function induced on simplices by a function $A \to B$.

Example 2.1 Suppose that $A = \{p, q, r, s, t\}$.

- Then $X = \{\{p, q\}, \{p\}, \{q\}, \{r\}\}$ is an abstract simplicial complex on A.
- Although $X' = \{\{q, p\}, \{p\}, \{q\}, \{r\}\}$ consists of the same *subsets* as X, we regard X and X' as being different simplicial complexes, since the ordering of the elements is different.
- $Y = \{\{p, q\}, \{p\}\}$ is not a simplicial complex, since it is missing the 0-simplex $\{q\}$.
- Both $Z = \{\{p, q\}, \{p, r\}, \{r, q\}, \{p\}, \{q\}, \{r\}\}$ and
- $W = \{\{p, q, r\}, \{p, q\}, \{p, r\}, \{r, q\}, \{p\}, \{q\}, \{r\}\}$ are abstract simplicial complexes.

Example 2.2 Suppose that $X = \{\{p, q\}, \{r, s\}, \{p\}, \{q\}, \{r\}, \{s\}\}$ and $Y = \{\{p, q\}, \{p\}, \{q\}\}$ are two simplicial complexes on $A = \{p, q, r, s\}$ and that $\phi : A \to A$ is a function given by

$$\phi(p) = p,\ \phi(q) = q,\ \phi(r) = p,\ \phi(s) = q.$$

This function induces the following simplicial map:

$$\{p, q\} \mapsto \{p, q\},\ \{p\} \mapsto \{p\},\ \{q\} \mapsto \{q\},\ \{r, s\} \mapsto \{p, q\},\ \{r\} \mapsto \{p\},\ \{s\} \mapsto \{q\}.$$

While very efficient, the definition of an abstract simplicial complex sometimes obscures topological information. It is therefore important to be able to *realize* a finite abstract simplicial complex X as a CW complex $|X|$ in which simplicial maps become cellular maps.

As shown in Fig. 2.12; each simplex in X corresponds to a cell in $|X|$, and is attached to each of its faces. We begin by constructing the realization of a single k-simplex.

Definition 2.10 The *standard k-simplex* is the closed subset of $\mathbb{R}^{k+1}$ given by

$$\Delta^k = \left\{ (x_1, \ldots x_{k+1}) : x_i \geq 0 \text{ for all } i, \text{ and } \sum_{i=1}^{k+1} x_i = 1 \right\}.$$

[3] The elements of A are need not be ordered themselves.

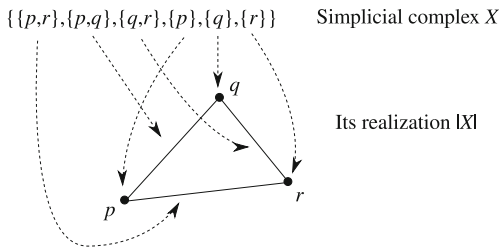

Fig. 2.12 A realization of an abstract simplicial complex

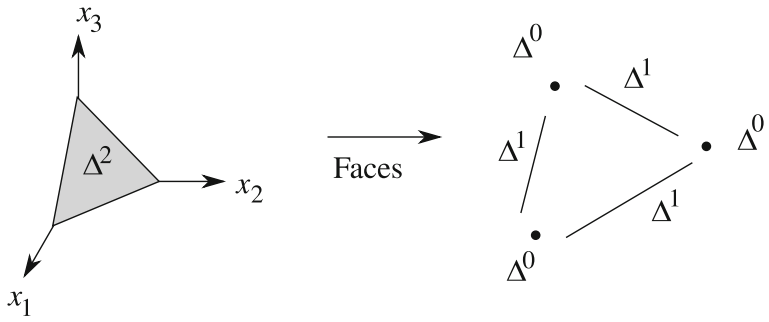

Fig. 2.13 The standard 2-simplex Δ^2 and its faces: Δ^0 are *points* and Δ^1 are *line* segments

Observe that Δ^k is the convex combination of the $k+1$ points whose coordinates are all 0 except one which is 1.

As can be seen in Fig. 2.13, intersecting the standard 2-simplex with the cardinal planes $x = 0$, $y = 0$, or $z = 0$ results in copies of the standard 1-simplex. This is generally true, since

$$\Delta^{k-1} = \Delta^k \cap \{(x_1, \ldots, x_{k-1}, 0)\}.$$

By permuting coordinates in the intersecting subspace, a total of $k+1$ copies of the standard $k-1$-simplex form the boundary of the standard k-simplex. It follows that intersecting Δ^k with a subspace formed by zeroing out n coordinates results in a copy of Δ^{k-n}.

Definition 2.11 Let X be an abstract simplicial complex on the finite set A. We define the *realization* $|X|$ to be the attachment construction given by $|X| = \left(\bigcup_\alpha \sigma_\alpha\right) / \sim$, where α is an index that ranges over all simplices in X and $\sigma_\alpha = \Delta^{k_\alpha}$ is a standard simplex of dimension k_α. We construct the attachment maps $\phi_\alpha : \partial\sigma_\alpha \to |X|^{k_\alpha - 1}$ so that each $(k_\alpha - 1)$-face of σ_α is mapped homeomorphically onto an element of $|X|^{k_\alpha - 1}$ as follows. Without loss of generality, suppose that σ_α corresponds to the k-simplex $\{1, 2, \ldots, k\}$. Consider ϕ_α restricted to the face $\{1, 2, \ldots, k-1\}$. Since X^k is an abstract simplicial complex, the simplex $\{1, 2, \ldots, k-1\}$ is also contained in X^k, and therefore corresponds to some other simplex σ_β. We therefore assign

$$\phi(\partial\sigma_\alpha) = \phi_\alpha(x_1, x_2, \ldots, x_{k-1}, 0) = (x_{j_1}, x_{j_2}, \ldots, x_{j_{k-1}}) \in \Delta^{k-1} = \sigma_\beta, \quad (2.1)$$

where $(1, \ldots, k-1) \mapsto (j_1, \ldots, j_{k-1})$ is a permutation. We note that assembling each ϕ_α from its restrictions to faces of σ_α (given by (2.1)) is automatically continuous, and it is therefore an attaching map.

Notice that since abstract simplicial complexes are closed under the operation of taking subsets, the restriction of ϕ_α to the boundary of the preimage of σ_β is automatically an attaching map. Further, each such ϕ_α is an embedding by construction. Hence $|X|$ is a cell complex.

Proposition 2.1 *If X is a CW complex whose attaching maps are embeddings, then there is an abstract simplicial complex Y for which $X \to |Y|$ is cellular homeomorphic to $|Y|$.*

In other words, for any CW complex X, one can find an abstract simplicial complex whose realization is a topological copy of X.

The statement is true if the attaching maps are not embeddings, though the proof is more difficult. We will not need that level of generality, however.

Proof We can proceed by induction on attaching maps. A 0-dimensional complex can be realized as a collection of vertices. If X^k can be realized as $|Y^k|$, consider attaching a $k+1$-cell e_α to X^k via the map ϕ. Suppose that $A_1 \cup \cdots \cup A_n$ is a decomposition such that all open cells in $\phi(e_\alpha)$ are one of the $\phi(A_i)$. If $n \leq k+2$, it is immediate (since ϕ is an embedding) that we merely need to add a single $k+1$-simplex to Y whose faces correspond to the $\phi(A_i)$. If $n > k+2$, then we merely need to subdivide e_α into simplices so that the union has n faces. (Figure 2.14 shows the 2-dimensional case.) □

Corollary 2.2 *If X is a noncompact cell complex, then it is cellular homeomorphic to $|Y|\backslash\{*\}$, the geometric realization of some abstract simplicial complex Y with a vertex removed. We will call $*$ the* external vertex.

Corollary 2.3 *If $f : X \to Y$ is a simplicial map, then f induces a cellular map $|f| : |X| \to |Y|$.*

2.2.2 Manifolds and Embeddings

While cell complexes have an overall notion of dimension, their dimension is not local. For example, the interior of an edge in a graph is homeomorphic to an open interval; a 1-dimensional space. Perhaps at a vertex attached to one or two edges, the dimension ought to be 1. But what about a vertex that is attached to three edges? It is unclear what dimension should be assigned to such a vertex. This situation does not arise in some applications, which constrains the allowable local topological structure in those cases. This leads naturally to the concept of a manifold.

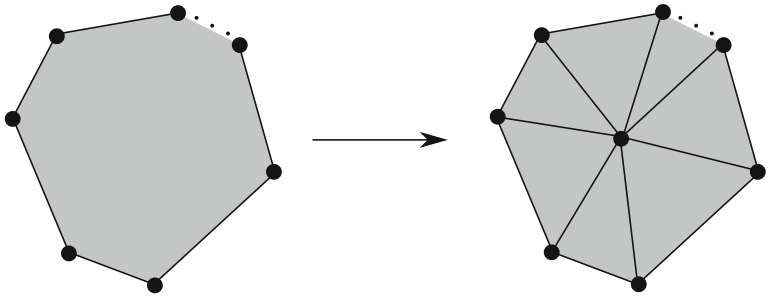

Fig. 2.14 Subdividing a 2-cell with n edges into simplices

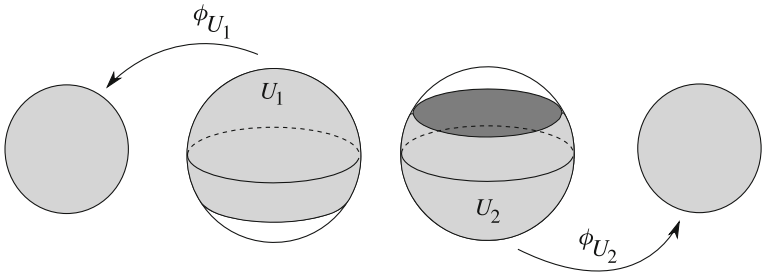

Fig. 2.15 An atlas with two charts on the 2-dimensional sphere S^2

Our study of CW complexes is intimately tied to the structure of manifolds; a compact manifold of a particular dimension can be given the structure of a CW complex of that dimension. (This is why we will assume manifolds to be Hausdorff[4] and paracompact,[5] since CW complexes satisfy both properties.) The most elegant proof of this fact involves the use of Morse theory. The author recommends several excellent texts on the subject, for instance Milnor (1963), Banyaga and Hurtubise (2004).

Definition 2.12 A *topological manifold* $(M, \mathscr{U})$ is a Hausdorff topological space M that is paracompact and has an *atlas*, a cover $\mathscr{U}$ of open sets in which each $U \in \mathscr{U}$ is homeomorphic to an open subset V of Euclidean space $\mathbb{R}^n$ through a map $\phi_U : U \to V$, called a *chart of* $\mathscr{U}$ (n may depend on U and is called the *local dimension* at U *and is written* $\dim U$).

Example 2.3 A sphere is an example of a manifold. The n-dimensional sphere is the following closed subset of $\mathbb{R}^{n+1}$

[4] A space is Hausdorff if every two distinct points are contained in disjoint open neighborhoods.

[5] A space is *paracompact* if every open cover $\mathscr{U}$ has a locally finite open refinement cover $\mathscr{V}$ in which each element $V \in \mathscr{V}$ is a subset of some $U \in \mathscr{U}$.

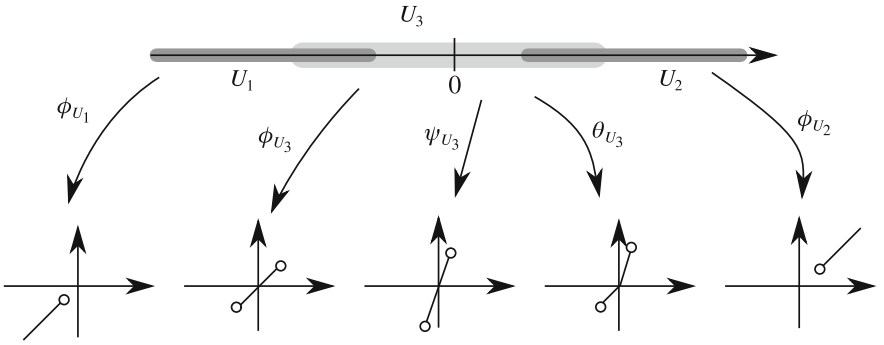

Fig. 2.16 Several different charts for $\mathbb{R}$

$$S^n = \left\{ (x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} x_i^2 = 1 \right\}.$$

One convenient atlas for the sphere consists of two overlapping hemispheres U_1, U_2, as shown in Fig. 2.15. (This atlas is analogous to the CW complex for the sphere shown in Fig. 2.8.) It's easiest to define ϕ_{U_1} and ϕ_{U_2} using the stereographic projection:

$$\phi_{U_1}(x, y, z) = \left(\frac{x}{1+z}, \frac{y}{1+z} \right) \text{ for } z > -1$$

and

$$\phi_{U_2}(x, y, z) = \left(\frac{x}{1-z}, \frac{y}{1-z} \right) \text{for } z < 1.$$

Definition 2.13 Suppose that $\mathscr{U}$ is an atlas of a topological manifold M. The *transition maps* of $\mathscr{U}$ are the functions $\phi_{U_1} \circ \phi_{U_2}^{-1}$ for $U_1, U_2 \in \mathscr{U}$, which are maps between open subsets of Euclidean space. If the transition maps have k continuous derivatives, then we say that $\mathscr{U}$ is a C^k *atlas*. If the transition maps have derivatives of all orders, then we say that $\mathscr{U}$ is a C^∞ *atlas* or a *smooth atlas*.

Exercise 2.3 Show that the transition maps $\phi_{U_1} \circ \phi_{U_2}^{-1}$ and $\phi_{U_2} \circ \phi_{U_1}^{-1}$ in Example 2.3 , which are functions $\mathbb{R}^2 \backslash \{(0,0)\} \to \mathbb{R}^2$, both have continuous derivatives away from the origin.

Definition 2.14 Two C^k atlases $\mathscr{U}$ and $\mathscr{V}$ are C^k *compatible* if their union $\mathscr{U} \cup \mathscr{V}$ is also a C^k atlas.

Example 2.4 This example shows several different atlases that can be put on the real line $\mathbb{R}$. Consider two open sets, $U_1 = (-\infty, -1)$, $U_2 = (1, \infty)$ with charts $\phi_{U_1} : U_1 \to \mathbb{R}$, $\phi_{U_2} : U_2 \to \mathbb{R}$ given by $\phi_{U_1}(x) = x$ and $\phi_{U_2}(x) = x$. We now consider another open set $U_3 = (-2, 2)$ with three different possible charts (see Fig. 2.16),

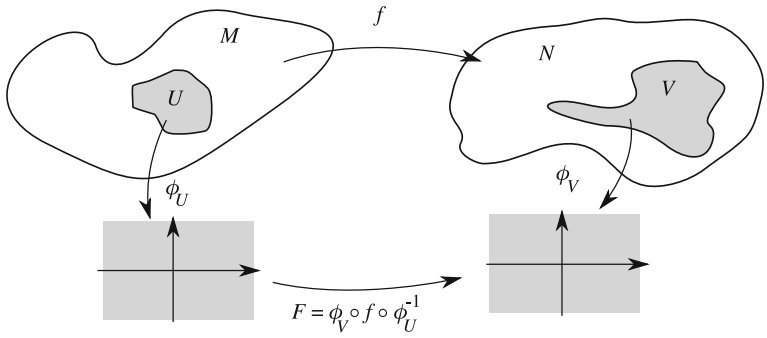

Fig. 2.17 A C^k function between two manifolds

$$\phi_{U_3}(x) = x, \ \psi_{U_3}(x) = 2x, \text{ and } \theta_{U_3}(x) = \begin{cases} x & \text{if } x < 0 \\ 2x & \text{if } x \geq 0. \end{cases}$$

The atlases $\{\phi_{U_1}, \phi_{U_2}, \phi_{U_3}\}$ and $\{\phi_{U_1}, \phi_{U_2}, \psi_{U_3}\}$ are smoothly compatible, since $\psi_{U_3} \circ \phi_{U_3}^{-1} = 2x$ is smooth. However, the atlases $\{\phi_{U_1}, \phi_{U_2}, \phi_{U_3}\}$ and $\{\phi_{U_1}, \phi_{U_2}, \theta_{U_3}\}$ are not compatible, since the transition map $\theta_{U_3} \circ \phi_{U_3}^{-1}$ has a discontinuous derivative.

Definition 2.15 Each C^k atlas $\mathscr{U}$ on a topological manifold M is compatible with a unique largest C^k atlas, in the sense of inclusion. We call this atlas a *maximal atlas*. A topological manifold $(M, \mathscr{U})$ in which $\mathscr{U}$ is a maximal C^k atlas is called a C^k *manifold*. If $\mathscr{U}$ is a smooth atlas, then we call $(M, \mathscr{U})$ a *smooth manifold*, or simply a *manifold*. Often we will abbreviate notation by stating "M is a smooth manifold with atlas $\mathscr{U}$" instead of stating "$(M, \mathscr{U})$ is a smooth manifold."

Continuous functions between manifolds that respect the manifold structure are given special status.

Definition 2.16 Suppose that $(M, \mathscr{U})$ and $(N, \mathscr{V})$ are manifolds and that $f : M \to N$ is a continuous function. We call f a C^k *map* or C^k *function* if for each $U \in \mathscr{U}$ and $V \in \mathscr{V}$ for which $f(U)$ intersects V, $F = \phi_V \circ f \circ \phi_U^{-1}$ restricted to $(\phi_U \circ f^{-1})(V)$ has k continuous derivatives. This is well-defined because F is a function between open subsets of Euclidean space, as shown in Fig. 2.17. The set of C^k maps between M and N is written $C^k(M, N)$. If F has derivatives of all orders, we say f is a *smooth map* in $C^\infty(M, N)$.

C^k manifolds are important in applications because they support a notion of calculus. In particular, we can define a notion of a derivative for C^k maps.

Definition 2.17 Suppose $f : M \to N$ is a C^k function between C^k manifolds with maximal atlases $\mathscr{U}$ and $\mathscr{V}$, respectively. Given $x \in U \in \mathscr{U}$ and $V \in \mathscr{V}$, the

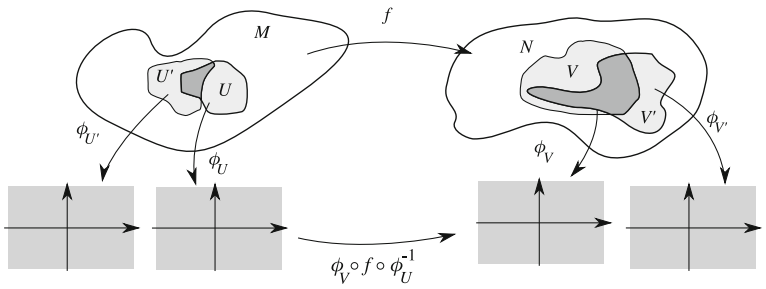

Fig. 2.18 Changes of charts on the domain and range of a C^k function between two manifolds

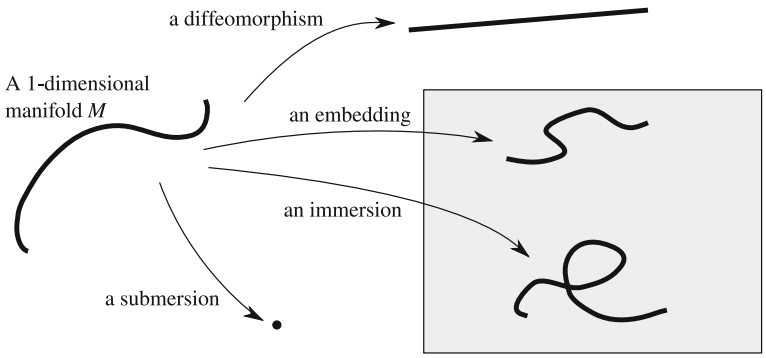

Fig. 2.19 Examples of different kinds of smooth maps of a 1-dimensional manifold

derivative of f at x is the Jacobian matrix of partial derivatives of $\phi_V \circ f \circ \phi_U^{-1}$. We will write $\mathrm{d}f_x$ for the derivative of f at x.

Lemma 2.1 *The rank of the derivative at a point $x \in M$ of a C^k function $f : M \to N$ between C^k manifolds is invariant under a change of charts.*

Proof Suppose that $\mathscr{U}$ and $\mathscr{V}$ are maximal C^k atlases of M and N, respectively. Observe that if x is also in $U' \in \mathscr{U}$ and $f(x)$ is also in V', then (see Fig. 2.18)

$$\begin{aligned}\phi_{V'} \circ f \circ \phi_{U'}^{-1} &= \phi_{V'} \circ \left(\phi_V^{-1} \circ \phi_V\right) \circ f \circ \left(\phi_U^{-1} \circ \phi_U\right) \circ \phi_{U'}^{-1} \\ &= \phi_{V'} \circ \phi_V^{-1} \circ \left(\phi_V \circ f \circ \phi_U^{-1}\right) \circ \phi_U \circ \phi_{U'}^{-1}\end{aligned}$$

The maps $\phi_U \circ \phi_{U'}^{-1}$ and $\phi_{V'} \circ \phi_V^{-1}$ are homeomorphisms between two Euclidean open sets and hence of full rank, a fact called the "invariance of dimension." Therefore $\phi_{V'} \circ f \circ \phi_{U'}^{-1}$ has the same rank as $\phi_V \circ f \circ \phi_U^{-1}$. □

Since dimension is an important local invariant of manifolds, maps that interact well with dimension have special names and properties as indicated in Fig. 2.19.

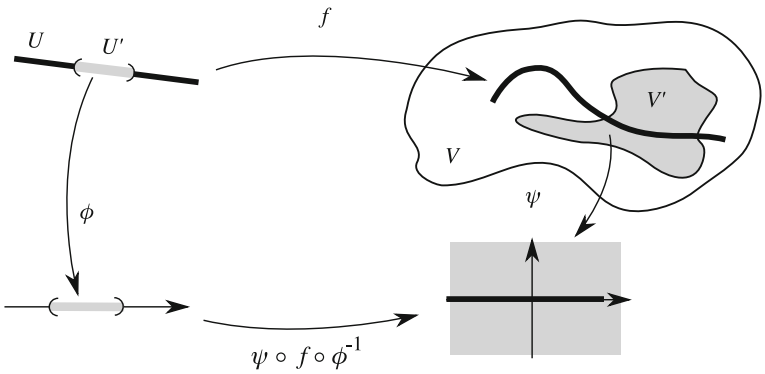

Fig. 2.20 A submanifold chart

Definition 2.18 Suppose $f : M \to N$ is a C^k function between C^k manifolds.

1. f is a *diffeomorphism* if it is a C^k homeomorphism with a C^k inverse. If f is a diffeomorphism, we say M and N are *diffeomorphic*.
2. If the derivative of f is injective at every point $x \in M$, then f is called an *immersion*.
3. An immersion that is also a homeomorphism onto its image is called an *embedding*.
4. If the derivative of f is surjective at every point $x \in M$, then f is called a *submersion*.

Written more simply, if a diffeomorphism exists between two manifolds, they are indistinguishable as manifolds. An embedding is a diffeomorphism into some ambient space. Immersions permit self-intersections. Submersions require the dimension of the domain to be greater than or equal to dimension of the range.

Definition 2.19 Suppose that N is a topological subspace of a manifold $(M, \mathscr{U})$. A *submanifold chart* for N is a chart $\phi : U \to \mathbb{R}^n$ of $\mathscr{U}$ in which $\phi(N \cap U)$ is the intersection of a linear subspace of $\mathbb{R}^n$ with $\phi(U)$. Explicitly, the *submanifold dimension* of N on U is the number k such that every point of the form $(x_1, \ldots, x_k, 0, \ldots, 0) \in \phi(U) \subseteq \mathbb{R}^n$ is the preimage of a point in N.

A submanifold $(N, \mathscr{V})$ of a manifold $(M, \mathscr{U})$ is a topological subspace of M which has an atlas $\mathscr{V}$ of submanifold charts.

Submanifold charts arise whenever there are C^k maps whose derivatives have constant rank.

Lemma 2.2 *(Loosely following (Lee 2003, Theorem 7.8)) Suppose $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^n$ are open and $f : U \to V$ is a C^1 map whose derivative has rank k at every point $x \in U$. For any $x \in U$, there exist charts $\phi : U' \to \mathbb{R}^m$ and $\psi : V' \to \mathbb{R}^n$ with $U' \subseteq U$ and $f(U) \subseteq V' \subseteq V$ such that (see Fig. 2.20)*

$$\psi \circ f \circ \phi^{-1}(x_1, \ldots, x_m) = (x_1, \ldots, x_k, 0, \ldots, 0).$$

The lemma is essentially a nonlinear form of the singular value decomposition, since the derivative of f has rank k at x we have

$$\mathrm{d}f_x = A^{-1} \begin{pmatrix} I_{k \times k} & 0_{k \times m-k} \\ 0_{n-k \times k} & 0_{n-k \times m-k} \end{pmatrix} B.$$

But the Lemma is actually stronger; this equality can be made to hold *everywhere* in the chart U'.

Proof Without loss of generality, we can assume that $x = 0$, and that $f(0) = 0$. For convenience, we will construct ϕ and ψ so that $\phi(0) = 0$ and $\psi(0) = 0$. Without loss of generality assume that by permuting dimensions in both the domain and range, we can write $f(x_1, \ldots, x_m) = (g(x_1, \ldots, x_m), h(x_1, \ldots, x_m))$, where $g : U \to \mathbb{R}^k$ has full rank and $h : U \to \mathbb{R}^{n-k}$. Given that we want $\psi \circ f \circ \phi^{-1}$ to be the identity on the first k dimensions, let us define

$$\phi(x_1, \ldots, x_k, \ldots, x_m) = (g(x_1, \ldots, x_k, \ldots, x_m), x_{k+1}, \ldots, x_m).$$

We claim that ϕ^{-1} is defined, at least if ϕ is restricted to some smaller open set $U' \subset U$, by the classical inverse function theorem. This follows because

$$\mathrm{d}\phi_0 = \begin{pmatrix} A_{k \times k} & B_{k \times m-k} \\ 0_{m-k \times k} & I_{m-k \times m-k} \end{pmatrix}$$

is nonsingular since $(A_{k \times k}\ B_{k \times m-k})$ is the matrix of partial derivatives of g, which is of full rank by assumption. Therefore we have that on U',

$$f \circ \phi^{-1}(x_1, \ldots, x_m) = (x_1, \ldots, x_k, c(x_1, \ldots, x_m)),$$

where $c : U' \to \mathbb{R}^{n-k}$ is a C^k function. Thus, we would like to define ψ to be

$$(y_1, \ldots, y_k, y_{k+1} - c_{k+1}(x_1, \ldots, x_m), \ldots, y_n - c_n(x_1, \ldots, x_m)).$$

Perversely, though, this means that ψ isn't a map to the subset of $\mathbb{R}^n$ we wanted. However, we happen to be lucky since c is independent of $x_{k+1}, \ldots, x_m$. Observe that

$$\mathrm{d}(f \circ \phi^{-1}) = \begin{pmatrix} I_{k \times k} & 0_{k \times m-k} \\ A_{n-k \times k} & B_{k \times m-k} \end{pmatrix}$$

must have column rank exactly equal to k, since ϕ is a C^k diffeomorphism by construction. Therefore, $B_{k \times m-k}$ is actually a zero matrix, proving our claim that c is independent of $x_{k+1}, \ldots, x_m$.

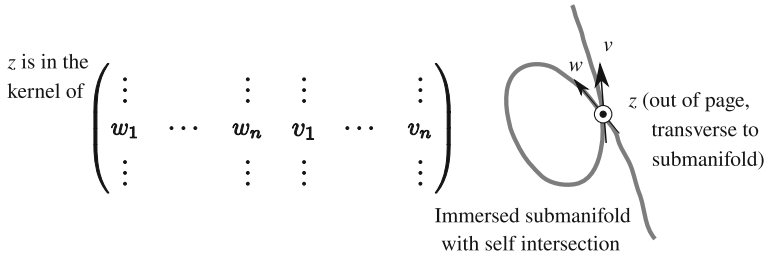

Fig. 2.21 Dimensional constraints in resolving self-intersections of an immersed manifold

Therefore, we have that $c(x_1, \ldots, x_k, x_{k+1}, \ldots, x_m) = C(x_1, \ldots, x_k)$, and so we may define

$$\psi(y_1, \ldots, y_n) = (y_1, \ldots, y_k, y_{k+1} - C_{k+1}(y_1, \ldots, y_k), \ldots, y_n - C_n(y_1, \ldots, y_k)).$$

□

This lemma facilitates the following proposition, that embeddings inject a copy of a manifold as a submanifold of another.

Proposition 2.2 *A C^k embedding $f : M \to N$ of one C^k manifold into another is also a C^k diffeomorphism onto its image $f(M)$, which is a submanifold of N.*

Proof Evidently, f is a homeomorphism when its range is restricted to $f(M)$. Since f is also an immersion, it is a map of constant (full) rank. Therefore, by the Lemma, we have that for every $x \in M$, there exist charts $\phi : U \subseteq M \to \mathbb{R}^m$ and $\psi : V \subseteq N \to \mathbb{R}^n$ so that $f(U) = f(M) \cap V$ (since f is a homeomorphism onto its image) and

$$\psi \circ f \circ \phi^{-1}(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0).$$

Hence ψ is a submanifold chart for $f(M)$ at $f(x)$. The inverse of f is C^k, since it is the identity with this choice of charts. □

The prototypical theorem about embeddings is the Whitney approximation theorem , which states that any compact manifold can be embedded in a Euclidean space of high enough dimension.

Theorem 2.1 *(Whitney approximation theorem (Whitney 1936, Theorem 2)) Suppose that $f : M \to \mathbb{R}^{2n+1}$ is a C^k map of a compact n-dimensional C^k manifold. Then for any $\epsilon > 0$, there exists a $g : M \to \mathbb{R}^{2n+1}$ such that $\|g(x)\| < \epsilon$ for every $x \in M$ and $f + g$ is an embedding. In this case, $\|v\|$ is the length of a vector v in $\mathbb{R}^{2n+1}$.*

The proof of Theorem 2.1 relies strongly on the (para) compactness of M, and a rather technical result called Sard's theorem. For an elementary proof, see Lee (2003).

More sophisticated versions involving transversality theory can be found in Golubitsky and Guillemin (1973). In contrast, an extremely elegant, elementary, and *constructive* approach is taken by Yomdin and Comte (2004) using something called "tame topology." In all cases, there are two essential aspects to the proof:

1. Showing that a C^k map can be perturbed[6] to a C^k immersion, and
2. Showing that a C^k immersion can be perturbed to a C^k embedding.

The dimensional limitation $\dim M \leq 2n+1$ comes from the following observation. Assume that our map f has a self intersection, that $f(x) = f(y)$ for $x \neq y$. At worst case, the derivatives at x and y span a subspace of dimension $2 \dim M$ as shown in Fig. 2.21. The matrix of derivatives

$$\begin{pmatrix} \mathrm{d}f_x & \mathrm{d}f_y \end{pmatrix}$$

is a matrix of size $2n+1 \times 2n$ and has rank at most $2n$. Thus, there exists at a vector $v \neq 0$ in the range of f that is not in the span of either derivative. We can move either x or y a small amount (bounded by ϵ) in the direction of v to remove the self intersection.

2.3 Case Study: Signal Manifolds for Localization, Tracking, and Navigation

When trying to "map out" an unknown environment, the locations of the measurements are of prime interest. *Remote imaging*, the process of inferring the location and properties of targets in a scene from measurements taken at a distance from the targets, has become an essential tool for cartographers. The measurements of a target need to be an unambiguous function of its position and other properties in order to be useful for locating and identifying it. Remote imaging requires *sources of illumination* or *transmitters* that provide signals that are scattered from targets in the scene and measured by *receivers*. We will assume that the transmitters and receivers are able to communicate amongst themselves without disrupting the collection process. This enables them to perform localization tasks using a smaller number of transmitters and receivers merely by moving them to different locations to form a "synthetic aperture."

Photographic techniques are often a preferred form of remote sensing due to their ease of interpretation. However, they suffer from many limitations due to weather, lighting conditions, and occlusions. For this reason, the use of lower frequencies both acoustically and electromagnetically has gained an indispensable role in cartography. Radio and sound waves can propagate long distances, which permits measurements to be taken far from the scene. In all of these cases, each location in the scene can be uniquely identified. For instance, positions in an optical image viewed by a

[6] A *perturbation* is a small change to a mathematical object.

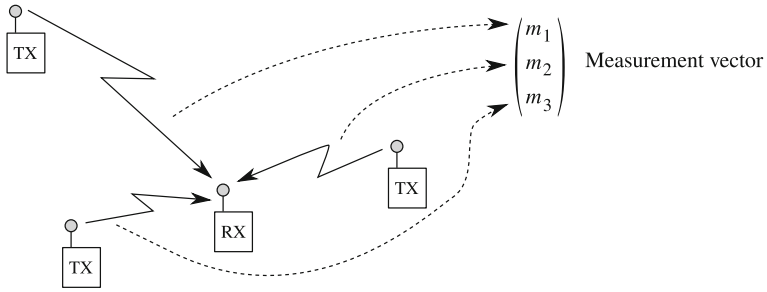

Fig. 2.22 Taking measurements of multiple signals at a receiver

camera or the human eye are discriminated by angle. (The accuracy of most optical sensors is usually quoted as an angular resolution.) Locations in a synthetic aperture radar image are discriminated by a unique distance to the sensor and associated Doppler frequency. If the signals in use do not travel in a straight line, it can seriously complicate interpretation, however. Magicians have often played with this ambiguity; the use of mirrors can result in an image for which the "easy" interpretation is incorrect and surprising. However, this ambiguity could be useful; reflective signals could permit imaging systems to "see around corners" (DARPA/STO 2010).

Remote imaging methods also play an important role in *navigation*, the method of inferring the location of a receiver in the scene. For instance, it is easy to see that measuring the distance from the receiver to three transmitters (in general position) whose locations are known suffices to determine the location of the receiver, if all are confined to move in a plane. What if the positions of the transmitters are not known, or the signal measurements do not correspond directly to distance? The Whitney approximation theorem provides a surprising answer: if the measurements of the signals depend smoothly on the receiver's location, then the receiver's location is completely determined using the signals from five (instead of three) transmitters. (Care must be exercised to ensure *transversality*—a generalization of general position—holds.)

Consider a setup as shown in Fig. 2.22. To each transmission link, the receiver takes a measurement, which is stored in a vector. To each receiver location, a different such vector is obtained, though we make the assumption that these vectors depend smoothly on the receiver location. Thus, all of the possible measurements can be encoded in a smooth map from the 2-dimensional plane (receiver locations) to a vector space whose dimension is the number of transmitters. The Whitney approximation theorem states that this map can be approximated by an injective map when the measurement dimension is more than twice the dimension of the plane, that is at least 5.

Realistic navigation systems based on this idea have additional complexity and redundancies. The most popular example is that of the Global Positioning System (GPS), in which transmitters are located on a constellation of 24 satellites (see Fig. 2.23) whose positions (as a function of time) are known very accurately. Each satellite transmits one of 12 coded signals; satellites with the same code have orbits

Fig. 2.23 A GPS satellite (image courtesy of US Air Force)

that place them on opposite sides of the earth. Therefore, every location on the earth is visible to at most 12 uniquely identifiable transmitters.

In addition to identity, the coded signals convey accurate timing information, which allows the receiver to synchronize its clock with the transmitter's clock. By measuring the apparent timing offsets between clocks associated to different transmitters, the receiver can measure its distance to each transmitter, thereby solving for its location. In order for this to succeed, the system of equations describing timing as a function of location must have a unique solution. This requires each location to have a unique collection of timing offsets.

Existing GPS systems also make the assumption that the signals propagate along straight lines, though they do account for propagation variations due to the Earth's gravitational field, relativity, and for average ionospheric conditions. This has an important implication: GPS does not work well in areas where reflective obstacles are present. For instance, consider the GPS track in Fig. 2.24 collected by the author near the University of Pennsylvania, in Philadelphia. The actual path was relatively straight (within a few feet), and remained on one side of the road. However, the GPS track shows erratic jumps in position, from one side of the road to the other. These jumps align fairly closely with the edges of nearby buildings. When he crossed 38th street (moving to the west), the buildings were much farther from the road and therefore diminished both the signal reflections and the GPS error. Because the track was not broken, we can infer that the author's position along this path was uniquely determined by the GPS signals. However, since there were occlusions and reflections, the GPS signals were no longer continuous (much less smooth) functions of his position. Therefore, the Whitney approximation theorem cannot be used in this

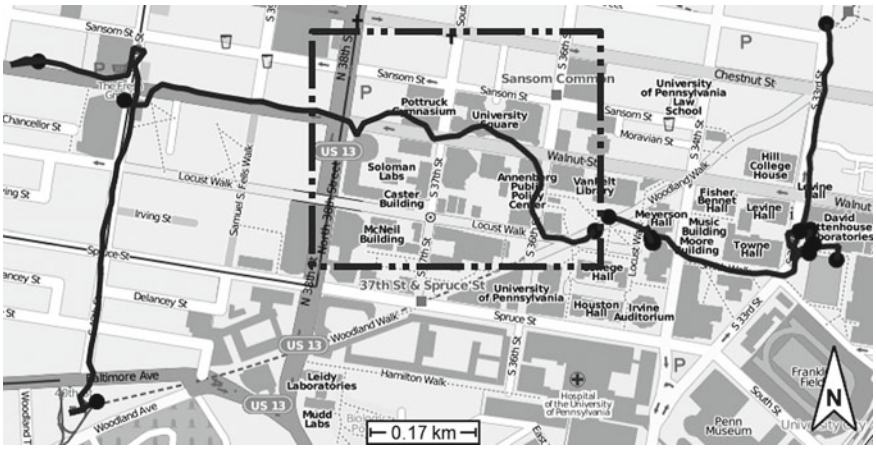

Fig. 2.24 An example GPS track collected by the author. East of 38th street, the buildings are close to the road and cause substantial error (*marked box*). West of 38th street, the buildings are far from the road so the error is reduced

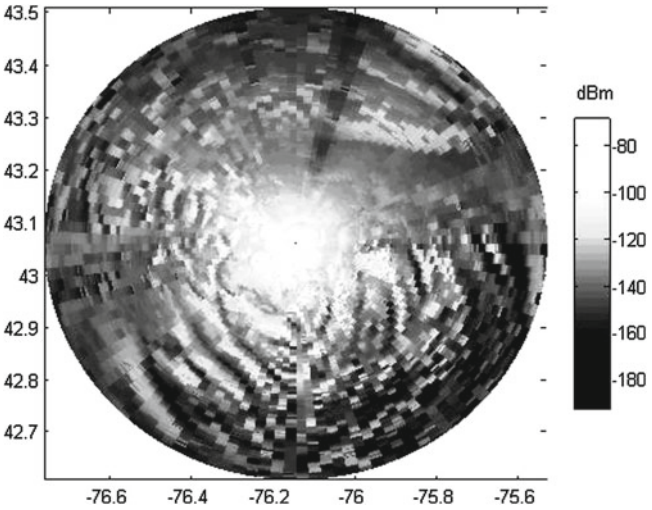

Fig. 2.25 The signal strength associated with a land-mobile radio communication tower located at the top of a hill varies dramatically with location. (Image produced using the model in Longley and Rice (1968) as implemented by the author, in collaboration with SRC, Inc.)

situation. However, a generalization of the Whitney approximation theorem called the Signal embedding theorem (Theorem 2.2), indicates that the receiver's location is still uniquely determined.

2.3.1 Signal Manifold Fingerprinting

Suppose that the propagation environment is a manifold M, and that there are n transmitters $T = \{t_1, \ldots, t_n\} \subset M$, and m receivers $R = \{r_1, \ldots, r_m\} \subset M$. Each transmitter emits a signal whose properties can vary over some portion of M. For instance, its signal strength, first time of signal arrival, polarization, antenna orientation, or bit error rate can vary dramatically according to various physical conditions as shown in Fig. 2.25. In order to provide substantial theoretical and practical flexibility, we take a very general approach to the received signal. Therefore, for each transmitter t_i, we assign

1. An open submanifold with compact closure $U_i \subseteq M$ describing the *coverage region* over which the signal can be reliably received,
2. A *signal manifold* S_i describing the possible parameters of the signal, for instance its signal strength, direction of arrival, or polarization, and
3. A smooth *propagation function* $P_i : U_i \to S_i$, which describes the received signal at every location in U_i.

Taken together, all of the transmitters generate a *transmission profile* that indicates all of the received signals at every given location. Formally, we extend each signal manifold with an additional point $\perp$, which indicates failure to receive that transmitter. Thus each propagation function extends to $P_i : M \to S_i \sqcup \perp$, which is likely to be discontinuous. We then collect all of these signal manifolds into a product, making the transmission profile a function $P : M \to \prod_{i=1}^{n} (S_i \sqcup \perp)$.

Unfortunately, the Whitney approximation theorem cannot be used on transmission profiles for two main reasons:

1. The overall collection of signals is discontinuous as a receiver crosses the boundary of a coverage region.
2. The signal perturbations associated to different transmitters are independent of one another.

However, a variant of the Whitney approximation theorem, called the signal embedding theorem (Theorem 2.2) does apply. In this more general setting, the result is somewhat weaker and does not imply that the transmission profile is an embedding in the sense of manifolds. However, it allows one to assert that an arbitrarily small perturbation of a transmission profile can be found which is injective and locally an embedding. The discontinuities of the perturbed transmission profile then result in a disconnected image in the signal manifold of a connected domain. In order to prove this result, we need a few preliminaries from transversality theory (Lee 2003, Golubitsky and Guillemin 1973).

Since the number of transmitters that each receiver can detect varies as a function of position, we require a bound on this number.

Definition 2.20 For a transmission profile P, the smallest number of non-$\perp$ components of $P(x)$ over all of M is called the *depth* of P, $\operatorname{dep} P$. (See the left frame of Fig. 2.26.)

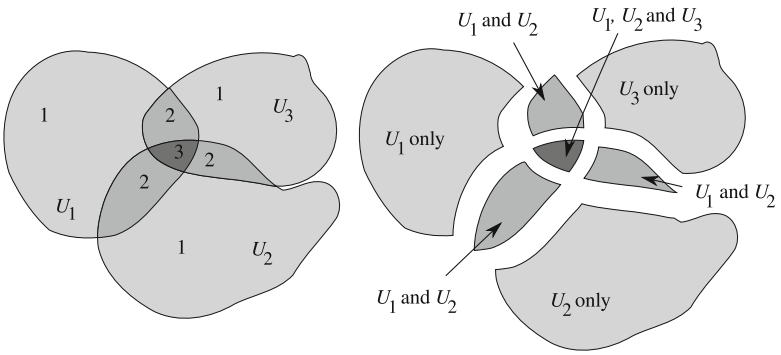

Fig. 2.26 Three coverage regions U_1, U_2, U_3 with mutual intersections. The resulting depth as a function of receiver location (*left*) and receiver locations partitioned according to transmitter coverage (*right*). The total depth in this example is 1

Definition 2.21 Suppose M and N are C^k manifolds and that $f : M \to \mathbb{R}^n$, $g : N \to \mathbb{R}^n$ are immersions. We say that f and g are *transverse at a point* x if their images intersect at x and the matrix of derivatives

$$\left(\mathrm{d}f_x \; \mathrm{d}g_x\right) \tag{2.2}$$

is of full rank.

Proposition 2.3 *If $f : M \to \mathbb{R}^n$ and $g : M \to \mathbb{R}^n$ are transverse at x, their images intersect in a submanifold of dimension $n - \dim M - \dim N$.*

Proof This result is an immediate consequence of the dimension theorem in linear algebra using two submanifold charts as constructed by Lemma 2.2. □

When the dimension $n - \dim M - \dim N < 0$, the method used to prove Theorem 2.1 (Whitney approximation) shows that a small perturbation of either map will result in no intersection at all. Stated more precisely, the set of maps in the space $C^k(M, \mathbb{R}^n)$ that do not intersect the image of g is open and dense (Golubitsky and Guillemin 1973). More generally, the set of maps in $C^k(M, \mathbb{R}^n)$ that are transverse to g is always open and dense when M is compact.

Theorem 2.2 *(Signal embedding theorem (Robinson and Ghrist 2012, Theorem 3)) There is an open and dense set of propagation functions for which the associated transmission profile is injective if* $2 \dim M < \mathrm{dep} P$.

Proof (sketch) Although the proof is fairly technical, the idea is simple: apply the Whitney approximation theorem on subsets of M illuminated by the same set of transmitters (right frame of Fig. 2.26). Points in different such subsets necessarily have different transmission profiles, and so cannot spoil the injectivity of the profile.

The technical complexity of the proof comes from showing that the transmission profile is injective on an open and dense set of propagation functions *independently*

of one another; we sketch the argument here. Suppose that the transmission profile is not injective, so that $P(x) = P(y)$. Take open neighborhoods V and W around x and y respectively, each contained entirely within the smallest intersections of coverage regions. Therefore, the transmission profile restricts to two smooth maps $f = P|_V$ and $g = P|_W$. The set of smooth maps from $V \to \mathbb{R}^{\mathrm{dep}P}$ that are transverse to g is open and dense, as noted before. On this set, the intersection of the image of such a map and the image of g will have dimension at most

$$\mathrm{dep}P - \dim V - \dim W = \mathrm{dep}P - 2\dim M,$$

since the image of the transmission profile has dimension at most $\mathrm{dep}P$. If this intersection dimension is less than zero, the set of maps whose image do not intersect the image of g will be open and dense. □

2.3.2 Multiple Target Detection and Localization

All forms of remote imaging rely on a unique signal response from each target in the scene; as a result, several kinds of systems are applications of the signal embedding theorem. As discussed previously, GPS is a popular navigation system that relies on straight-line distances to satellite transmitters. Consider the satellite transmitters with locations $t_1, \ldots, t_{12}$, and the receiver with location r. If there are no reflections and the transmitters are synchronized to one another, the time of arrival of each signal at r is $\tau_i = d(t_i, r)/c$ where c is the speed of light and $d(\cdot, \cdot)$ computes the distance between two points (corrected for relativistic effects). Since the GPS receiver does not have an accurate absolute time reference, it computes differences between the first time of arrival and all the others. Therefore, the transmission profile is a smooth map that takes $r \mapsto (\tau_2 - \tau_1, \ldots, \tau_{12} - \tau_1)$.

The standard algorithm assumes that the signals propagate along geodesics, and this map can be inverted explicitly. However, if six satellites are always in view, then the transmission profile has depth 5. According to the signal embedding theorem, the transmission profile will be injective on the 2-dimensional surface of the earth, *even if some of the signals have been reflected or occluded*. If instead, seven satellites are always in view, then the transmission profile can be used to discriminate between receivers with the same latitude and longitude but different heights.

One could also take the transmission profile to be $(r, t_1) \mapsto (\tau_2 - \tau_1, \ldots, \tau_{12} - \tau_1)$, which encodes the desire to solve for synchronization as well as location. In this case, eight satellites could be used to solve for 3-dimensional position and absolute time, again even if the signals have been reflected off buildings. At present, no GPS receivers use this idea, as it would apparently require tabulating or computing the expected reflections from known buildings. However, the signal embedding theorem is important because it places *lower* bounds on what is possible.

Taking the GPS idea further, one can imagine a radar system called a *moving target indicator*, which tracks a number of targets (Stimson 1998). In this case,

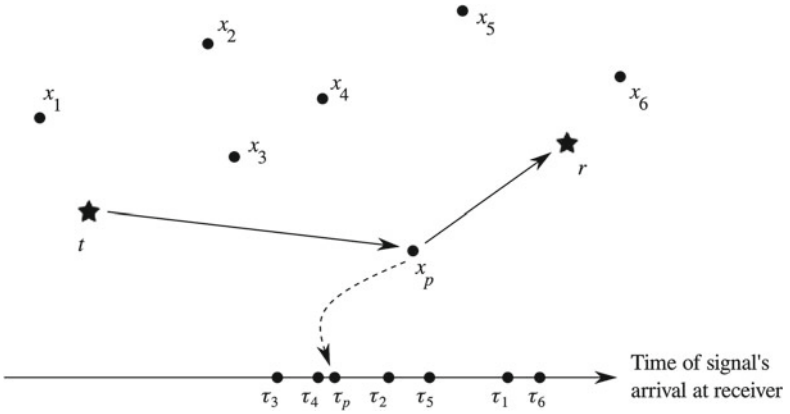

Fig. 2.27 Multiple targets being tracked by a transmitter and receiver

suppose there is a single transmitter t, receiver r, and a number of targets $x_1, \ldots, x_p$ as shown in Fig. 2.27. We can safely assume that the transmitter and receiver have known locations and are synchronized, but that the scene may have complicated reflective and obstructive geometry. Assume that the first time a signal traverses the path $t \to x_i \to r$ is given by the propagation function $\tau_i(x)$, which is smooth on some open set of positions for the target $U_i \subseteq M$. (Different targets may be more or less visible, so τ_i and τ_j may be different.) Outside of U_i, τ_i takes the value $\perp$.

We can then assemble a transmission profile from the τ_i, which has a depth of at most p. However, since there are p targets, the domain of the transmission profile has dimension $p \dim M$, which is too large for the signal embedding theorem. In this case, the signal embedding theorem is indicating that the problem is underdetermined. We can supply additional information in a variety of ways. For one, we can augment τ_i with an additional dimension by measuring the signal strength of the response. However, experience with radar and sonar systems suggests that this is somewhat unreliable, due to substantial variations in scattering cross section (Stimson 1998).

Instead, the usual way to solve the problem is to add multiple transmitters or receivers (or both) to supply additional measurements. If we add additional transmitter-receiver pairs for a total of q pairs and construct them so that they do not interfere, then the depth could be as large as pq. Thus, in the best case scenario $q > 2 \dim M$ would suffice to uniquely solve for each target's location, again without constraints on the geometry.

Having multiple transmitters and receivers is often expensive, so many remote sensing systems utilize moving sensors to form a synthetic apertures, see Jakowatz et al. (1996), Carrara et al. (1995). (See Fig. 2.28 for an example image product.) They usually rely on the choice of a coded waveform that permits the correlation of temporal characteristics against other properties of the scene. In this way, if a unique location can be extracted from the timing alone, one can then use the location to extract other properties. Specifically, consider the case of a moving transmitter $t(\tau)$

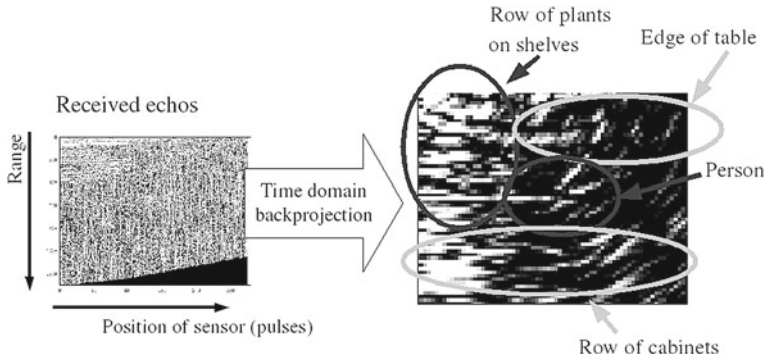

Fig. 2.28 An example synthetic aperture sonar image of the author's kitchen

and receiver $r(\tau)$. If M has a metric d on it, then we can write the range (distance from transmitter, to target, to receiver) and Doppler frequency (rate of change of range) explicitly. For location $x \in M$ in the scene, one obtains a range (from arrival times)

$$R(\tau, x) = d(t(\tau), x) + d(x, r(\tau)),$$

and a range rate (from the Doppler frequency offset)

$$D(\tau, x) = \frac{\partial}{\partial \tau} R(\tau, x).$$

Observe that the map $x \mapsto (R(\tau_1, x), D(\tau_1, x), \ldots, R(\tau_p, x), D(\tau_p, x))$ is a signal profile whose depth is $2p$. Hence, if $p > \dim M$, the signal embedding theorem ensures that we can discriminate locations (and hence material properties at those locations) based on timing alone.

2.4 Open Questions

1. The signal embedding theorem, while helpful from a feasibility point of view, is unfortunately nonconstructive. If it states that the transmission profile is injective, it provides no insight about how to implement a system to exploit the profile. The obvious implementations of the applications we have outlined require substantial tabulation of previous measurements, which is probably not ideal. At present, no constructive algorithms exploit the transmission profile directly.
2. Within each coverage region, the propagation function is smooth and perhaps therefore somewhat uninformative. Traditional methods focus on propagation functions, because they can be completely characterized. This is a weakness if there are many reflections and occlusions. On the other hand, the boundaries of

the coverage regions are both well-defined and lower dimensional. Crossing a shadow boundary therefore conveys substantial location information. Can one organize the shadow boundary information more efficiently or effectively than the individual propagation functions?

3. If there is a sufficiently dense set of reflective targets, these coverage region boundaries will be rather prevalent. However, the coverage regions are now dependent on locations of scatterers as well as the locations of the transmitters and receivers, and several coverage regions may arise from each illumination. What geometric bounds arise are there on the prevalence of shadows?
4. The collection of intersections of coverage regions forms a cell complex, in which the cells are labeled with which transmitters are detectable. Are there constraints on how typical faces of these cells could be labeled? Are there statistical or asymptotic properties that should be expected when the number of cells is large?
5. Traditional image formation methods require careful synchronization between transmitter and receiver, as assumed in our discussion above. If relaxed, one can construct a transmission profile to solve for the unknown time offsets. However, it is unclear what algorithms are possible in this context. Specifically, is there a general algorithmic formulation of synthetic aperture image formation that is not dependent on knowledge of the collection geometry and timing?

References

Banyaga A, Hurtubise D (2004) Morse homology. Springer, Berlin

Carrara W, Goodman R, Majewski R (1995) Spotlight synthetic aperture radar: signal processing algorithms. Artech House, Norwood

DARPA/STO: MER data collection review (2010). http://www.darpa.mil/STO/solicitations/baa09-01/presentations/MER_Data_Review.pdf

Golubitsky M, Guillemin V (1973) Stable mappings and their singularities. Springer, New York

Jakowatz C, Wahl D, Eichel P, Ghiglia D, Thompson P (1996) Spotlight-mode synthetic aperture radar: a signal processing approach. Kluwer Academic Publishers, Dordrecht

Kaczynski T, Mischaikow K, Mrozek M (2004) Computational homology. Springer, New York

Lee J (2003) Smooth manifolds. Springer, New York

Longley A, Rice P (1968) Prediction of tropospheric radio transmission loss over irregular terrain. A computer method. Technical Report. Boulder, CO

Milnor J (1963) Morse theory. Princeton University Press, Princeton

Robinson M, Ghrist R (2012) Topological localization via signals of opportunity. IEEE Trans Sig Proc 60(5):2362–2373

Steen L, Seebach JA (1978) Counterexamples in topology. Dover, New York

Stimson GW (1998) Airborne radar. SciTech, Raleigh

Whitney H (1936) Differentiable manifolds. Annal Math Second Ser 36(3):645–680

Yomdin Y, Comte G (2004) Tame geometry with application in smooth analysis. Springer, Heidelberg

http://www.springer.com/978-3-642-36103-6

Topological Signal Processing
Robinson, M.
2014, XVI, 208 p. 134 illus., Hardcover
ISBN: 978-3-642-36103-6