Subdivision of Maps of Digital Images

Gregory Lupton1 · John Oprea1 · Nicholas A. Scoville2

Received: 25 June 2019 / Revised: 19 July 2021 / Accepted: 27 July 2021 / Published online: 7 February 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
A digital image is a finite set of integer lattice points in an ambient Euclidean space together with a suitable adjacency relation between points. Subdivision, which is a process of enlarging a digital image in the photographic sense, provides a basic tool for operating with digital images. But given a map of digital images, there is as yet no general way to define a map of their subdivisions that might reasonably be called a subdivision of the map. In this paper, we construct such maps of subdivisions when the map of digital images has a 1- or 2-dimensional domain. From our constructions we deduce path covering and homotopy covering results that play a role in our development of the digital fundamental group.

Keywords Digital image · Digital topology · Subdivision · Subdivision of a map · Digital fundamental group

Mathematics Subject Classification 55P99 · 54A40 · 68U10
1 Introduction

In digital topology, the basic object of interest is a digital image: a finite set of integer lattice points in an ambient Euclidean space with a suitable adjacency relation between points. This is an abstraction of an actual digital image which consists of pixels (in the plane, or higher dimensional analogues of such).

There is an extensive literature with many results that use ideas from topology in this setting (e.g. [4,9,15]). In many instances, however, notions from topology have been translated directly into the digital setting in a way that results in digital versions of topological notions that are very rigid and hence have limited applicability. In contrast to this existing literature, in [13,14] and this paper, we have started to build a more general “digital homotopy theory” that brings the full strength of homotopy theory to the digital setting. In our approach, we aim to use less rigid constructions, with a view towards broad applicability and greater depth of development. A key ingredient in such an approach is subdivision. However, the behavior of maps with respect to subdivision is not well-understood. In this paper, we establish fundamental results about subdivision of maps of digital images with 1-dimensional (1D) and 2-dimensional (2D) domains. The utility of these results is indicated in [13], in which we define a digital fundamental group and show that it is an invariant of subdivision-homotopy equivalence, which is a concept of “sameness” for spaces that is much less rigid than the notion of homotopy equivalence that is commonly used in digital topology. Our results of [13,14], both in the basic constructions and in the developments, emphasize subdivision as a basic feature, whereas in those of [4] and many other articles in the digital topology literature, subdivision plays a background role at most. Our results here on subdivision of maps also allow us to define invariants of 2D digital images such as Lusternik–Schnirelmann category in a way that is much less rigid than previously done (e.g. as in [3]). In general, our results work towards establishing “subdivision versions” of the usual invariants. Our motivating point of view is that one should incorporate subdivision at a basic level, rather than directly translate a definition or construction from the topological to the digital setting. Incorporating subdivision results in digital invariants whose behavior more closely follows that of their topological counterparts, when compared to the commonly used digital invariants that do not incorporate subdivision. To do this generally, however, requires a fuller understanding of the behavior of maps with respect to subdivision—maps with domains of arbitrary dimension.

The paper is organized as follows. In Sect. 2 we review standard definitions and terminology, and set our conventions (especially with regard to adjacency). In Sect. 3 we give a thorough discussion of subdivision of digital images and maps of digital images. We show how subdivision may be broken down into a succession of partial subdivisions (Corollary 3.9). Several figures are included that serve to indicate the basic ideas and concerns. The main question, illustrated through examples, is how— or even whether—a map of digital images induces one on subdivisions. In Sect. 4 we resolve this question for maps of digital images whose domain is an interval, namely paths and loops in a digital image (of any dimension). In Sect. 5 we do likewise for maps whose domain is a 2D digital image. In each case, we construct a canonical map of subdivisions from a given map of digital images. The main results are Theorem 4.1
and Theorem 5.9. A brief indication of the way in which our results here may be applied is given in Sect. 6. But applications of and developments from these results appear elsewhere. There, we also indicate how our results here on subdivision of maps lay the groundwork for future developments.

2 Basic Notions: Adjacency, Continuity, Products

In this paper, a digital image (of dimension n) X means a finite subset \( X \subseteq \mathbb{Z}^n \) of the integral lattice in some n-dimensional Euclidean space, together with a particular adjacency relation inherited from that of \( \mathbb{Z}^n \). Namely, two points \( x = (x_1, \ldots, x_n) \in \mathbb{Z}^n \) and \( y = (y_1, \ldots, y_n) \in \mathbb{Z}^n \) are adjacent if their coordinates satisfy \( |x_i - y_i| \leq 1 \) for each \( i = 1, \ldots, n \).

**Remark 2.1** In the literature, it is common to allow for various choices of adjacency. For example, a planar digital image is a subset of \( \mathbb{Z}^2 \) with either “4-adjacency” or “8-adjacency” (see, e.g. [4, Sect. 2]). However, in this paper, we always assume (a subset of) \( \mathbb{Z}^n \) has the highest degree of adjacency possible (8-adjacency in \( \mathbb{Z}^2 \), 26-adjacency in \( \mathbb{Z}^3 \), etc.). In fact, there is a philosophical reason for our fixed choice of adjacency relation: It is effectively forced on us by the considerations of Definition 2.3 and Example 2.5 below.

If \( x, y \in X \subseteq \mathbb{Z}^n \), we write \( x \sim_X y \) to denote that \( x \) and \( y \) are adjacent in \( X \). For digital images \( X \subseteq \mathbb{Z}^n \) and \( Y \subseteq \mathbb{Z}^m \), a function \( f : X \rightarrow Y \) is continuous if \( f(x) \sim_Y f(y) \) whenever \( x \sim_X y \). By a map of digital images, we mean a continuous function. Occasionally, we may encounter a non-continuous function of digital images. But, mostly, we deal with maps—continuous functions—of digital images. The composition of maps \( f : X \rightarrow Y \) and \( g : Y \rightarrow Z \) gives a (continuous) map \( g \circ f : X \rightarrow Z \), as is easily checked from the definitions.

An isomorphism of digital images is a continuous bijection \( f : X \rightarrow Y \) that admits a continuous inverse \( g : Y \rightarrow X \), so that we have \( f \circ g = \text{id}_Y \) and \( g \circ f = \text{id}_X \), and \( g \) is also bijective. If \( f : X \rightarrow Y \) is an isomorphism of digital images, then we say that \( X \) and \( Y \) are isomorphic digital images, and write \( X \cong Y \).

**Example 2.2** We use the notation \( I_N \) for the digital interval of length \( N \), namely \( I_N \subseteq \mathbb{Z} \) consists of the integers from 0 to \( N \) in \( \mathbb{Z} \), and consecutive integers are adjacent. Thus, we have \( I_1 = [0, 1] = \{0, 1\} \), \( I_2 = [0, 2] = \{0, 1, 2\} \), and so on. Occasionally, we may use \( I_0 \) to denote the singleton point \( \{0\} \subseteq \mathbb{Z} \). As an example in \( \mathbb{Z}^2 \), consider what we call the Diamond, \( D = \{(1, 0), (0, 1), (-1, 0), (0, -1)\} \), which may be viewed as a digital version of a circle (called a digital simple closed curve in the literature [4]). Note that pairs of vertices all of whose coordinates differ by 1, such as \((1, 0) \) and \((0, 1) \) here, are adjacent according to our definition. Otherwise, \( D \) would be disconnected. In Fig. 1 we have included the axes (dashed) and also indicated adjacencies (solid) in the style of a graph. Note, though, that we have no choice as to which points are adjacent: this is determined by position, or coordinates, and we do not choose to add or remove edges here. As an example in \( \mathbb{Z}^3 \), we have \( S = \{(1, 0, 0), (0, 1, 0), (-1, 0, 0), (0, -1, 0), (0, 0, 1), (0, 0, -1)\} \).
(the vertices of an octahedron, with adjacencies corresponding to the edges of the octahedron). This may be viewed as a digital 2-sphere, and the pattern emerging here may be continued to a digital $n$-sphere in $\mathbb{Z}^{n+1}$ with $2n + 2$ vertices. The map $f : I_2 \to I_1$ given by $f(0) = 0$, $f(1) = 0$, and $f(2) = 1$ is continuous, but the function $g : I_1 \to I_2$ given by $g(0) = 0$, $g(1) = 2$ is not: we cannot “stretch” an interval to a longer one. Likewise, suppose we enlarge $D$ to the larger digital simple closed curve $C = \{(2, 0), (1, 1), (0, 2), (-1, 1), (-2, 0), (-1, -1), (0, -2), (1, -1)\}$ (see Fig. 1). Then we cannot “stretch” the smaller circle around the larger by a map, in the sense that there are too few points in $D$ to make a complete (continuous) circuit of $C$. Digital simple closed curves and $n$-spheres of this kind, as well as disks, appear frequently in the literature and from a variety of points of view. See, for example, [1,12] and the references contained therein. A common approach—although not one we use—is to describe and analyze circles and disks in terms of a suitable distance function (see [11,16]). For instance our $D$ and $C$ here may be described as circles of radius 1 and 2, respectively, using the distance function $d_1$ discussed in [16] (neighbors of a point in our sense here may be described as points of distance $\leq 1$ from the point using the distance function $d_2$ of [16]). There is also considerable literature on the algorithmic and effective generation of digital circles and disks (see [2]). For another point of view, circles and spheres are discussed in [10] in a more general graph-theoretic setting.

A comment in the preceding example points to the main motivation for the results of this paper. Whereas homotopy is not the main focus of this paper (the notion is reviewed here in Sect. 6), our results here are motivated by wanting to relax the notion of homotopy equivalence commonly used in digital topology. We can give the basic idea informally, as follows. Because we cannot “stretch” the smaller circle around the larger one, there are too few maps from $D$ to $C$ of Example 2.2 for these digital simple closed curves to be homotopy equivalent, in the sense commonly used in digital topology. This assertion is justified in detail in [13, Exam. 3.32]. But from
a (topological) homotopy point of view, it seems reasonable to view $D$ and $C$—more generally, digital simple closed curves of different sizes—as being equivalent. In [13, Exam. 3.32], we develop a notion of subdivision-homotopy equivalence of digital images, which is a notion of “sameness” of digital images that combines subdivision with homotopy equivalence, and which is a less rigid notion of “sameness” than digital homotopy equivalence. Indeed, it turns out that $D$ and $C$ are subdivision-homotopy equivalent, but not homotopy equivalent [13, Exam. 3.32].

**Definition 2.3** (digital products) The product of digital images $X \subseteq \mathbb{Z}^m$ and $Y \subseteq \mathbb{Z}^n$ is the Cartesian product of sets $X \times Y \subseteq \mathbb{Z}^m \times \mathbb{Z}^n \cong \mathbb{Z}^{m+n}$ with the adjacency relation $(x, y) \sim_{X \times Y} (x', y')$ when $x \sim_X x'$ and $y \sim_Y y'$.

In fact, this is tantamount to our assumption that $\mathbb{Z}^n$, and any digital image in it, has the highest degree of adjacency possible, with the isomorphisms $\mathbb{Z}^n \cong \mathbb{Z}^1 \times \mathbb{Z}^{n-1}$ for $r = 1, \ldots, n - 1$. Note that some authors in the literature use a different adjacency relation on the product: the graph product, whereby $(x, y)$ is adjacent to $(x', y')$ if $x = x'$ and $y \sim_Y y'$, or $x \sim_X x'$ and $y = y'$. The notion we use is sometimes called the strong product, in a graph theory setting. Our definition of (adjacency on) the product means that it is the categorical product, in the category of (finite) digital images and digitally continuous maps. This point is explained in the following statement.

**Lemma 2.4** For digital images $X \subseteq \mathbb{Z}^m$ and $Y \subseteq \mathbb{Z}^n$, the projections onto either factor $p_1: X \times Y \to X$ and $p_2: X \times Y \to Y$ are continuous. Suppose given maps of digital images $f: A \to X$ and $g: A \to Y$. Then there is a unique map, which we write $(f, g): A \to X \times Y$ that satisfies $p_1 \circ (f, g) = f$ and $p_2 \circ (f, g) = g$.

**Proof** The first assertion follows immediately from the definitions. The map $(f, g)$ is defined as $(f, g)(a) = (f(a), g(a))$. It is immediate from the definitions that this map is continuous. This is evidently the unique map with the suitable coordinate functions. \hfill \Box

**Example 2.5** For $X \subseteq \mathbb{Z}^n$ a digital image, the diagonal map

$$\Delta: X \to X \times X \subseteq \mathbb{Z}^n \times \mathbb{Z}^n \cong \mathbb{Z}^{2n}$$

is defined as $\Delta(x) = (x, x)$ for each $x \in X$. Suppose we have $X = I_1 \subseteq \mathbb{Z}$, with $\Delta: I_1 \to I_1 \times I_1$. Since $0 \sim_X 1$, we need $(0, 0) \sim_{X \times X} (1, 1)$ if the diagonal is to be continuous, which of course we do have with our conventions.

Because of the rectangular nature of the digital setting that we use, it is often convenient to consider the product of maps, as follows.

**Definition 2.6** Given functions of digital images $f_i: X_i \to Y_i$ for $i = 1, \ldots, n$, we define the product function

$$f_1 \times \cdots \times f_n: X_1 \times \cdots \times X_n \to Y_1 \times \cdots \times Y_n$$

as $(f_1 \times \cdots \times f_n)(x_1, \ldots, x_n) = (f_1(x_1), \ldots, f_n(x_n))$.  

\[ Springer\]
Lemma 2.7 Given continuous maps of digital images \( f_i : X_i \to Y_i \) for \( i = 1, \ldots, n \), their product \( f_1 \times \cdots \times f_n \) is a (continuous) map.

Proof This follows directly from the definitions. \( \square \)

We will make use of the product of maps towards the end of the following section and in the sequel. This gives another reason for why we want the product of digital images to be defined as in Definition 2.3.

Another construction that we use—mainly in the context of proofs—is the union of digital images. For two digital images \( X, Y \subseteq \mathbb{Z}^n \) in the same ambient \( \mathbb{Z}^n \) as each other, their union \( X \cup Y \subseteq \mathbb{Z}^n \) is the ordinary set-theoretic union of the two sets \( X \) and \( Y \) with the adjacency inherited from \( \mathbb{Z}^n \). Occasionally we treat disjoint unions of digital images and use the variant notation “⊔” (or its displayed counterpart “\( \sqcup \)”) to emphasize the disjointness. In either case the union is the same construction. The main difference—so far as this paper is concerned—is that when assembling a (continuous) map from maps defined on subsets of a digital image, it is not necessary to check agreement on overlaps if the subsets are disjoint from each other.

3 Subdivision

The notion of subdivision of a digital image plays a fundamental role in our development of ideas in the digital setting, and is a main focus of this paper.

Definition 3.1 Suppose that \( X \subseteq \mathbb{Z}^n \) is an \( n \)-dimensional digital image. For each integer \( k \geq 2 \), we have the \( k \)-subdivision of \( X \), which is an auxiliary (to \( X \)) \( n \)-dimensional digital image denoted by \( S(X, k) \subseteq \mathbb{Z}^n \), together with a canonical map or standard projection

\[
\rho_k : S(X, k) \to X
\]

that is continuous in our digital sense. For a real number \( x \), denote by \( \lfloor x \rfloor \) the greatest integer less-than-or-equal-to \( x \) (the integer floor of \( x \)). First, make the \( \mathbb{Z}[1/k] \)-lattice in \( \mathbb{R}^n \), namely, those points with coordinates each of which is \( z/k \) for some integer \( z \), and then set

\[
S'(X, k) = \left\{ (x_1, \ldots, x_n) \in \left( \mathbb{Z} \left[ \frac{1}{k} \right] \right)^n \mid (\lfloor x_1 \rfloor, \ldots, \lfloor x_n \rfloor) \in X \right\}.
\]

Then set

\[
S(X, k) = \{ (kx_1, \ldots, kx_n) \in \mathbb{Z}^n \mid (x_1, \ldots, x_n) \in S'(X, k) \}.
\]

The map \( \rho_k \) is given by \( \rho_k((y_1, \ldots, y_n)) = (\lfloor y_1/k \rfloor, \ldots, \lfloor y_n/k \rfloor) \), and one checks that this map is continuous.

For \( x \in X \) an individual point, we write \( S(x, k) \subseteq S(X, k) \) for the points \( y \in S(X, k) \mid \rho_k(y) = x \). If \( x = (x_1, \ldots, x_n) \) is a point in an \( n \)-dimensional digital image, then we may describe this set in general as

\( \square \) Springer
\[ S(x, k) = \{(kx_1 + r_1, \ldots, kx_n + r_n) | 0 \leq r_i \leq k - 1\}. \quad (1) \]

That is, for each \( x \in X \), \( S(x, k) \) is an \( n \)-dimensional cubical lattice in \( \mathbb{Z}^n \) with each side of the cubical lattice containing \( k \) points. Notice that the result of subdivision therefore depends on the ambient space of the digital image.

Occasionally, it may be convenient to extend Definition 3.1 to include \( k = 1 \), in which case we use the notational convention that \( S(X, 1) = X \), and \( \rho_1 : S(X, 1) \to X \) is just the identity map of \( X \).

**Example 3.2** Generally, subdivision of an interval \( I_N \subseteq \mathbb{Z} \) gives a longer interval: We have \( S(I_N, k) = I_{Nk+k-1} \subseteq \mathbb{Z} \). Suppose that we have \( X = I_2 = [0, 2] \subseteq \mathbb{Z}^2 \).

Then we have \( S(I_2, 2) = I_5 = [0, 5] \subseteq \mathbb{Z} \), and \( \rho_2 : S(I_2, 2) \to X \) is given by \( \rho_2(0) = \rho_2(1) = 0, \rho_2(2) = \rho_2(3) = 1, \) and \( \rho_2(4) = \rho_2(5) = 2 \). In Fig. 2, we indicate the way in which, for the same interval \( I_2 \), the projection \( \rho_3 : S(I_2, 3) \to I_2 \) aggregates points in the subdivided interval to map them back to the original. We also note here that \( S(I_0, k) = S((0), k) = I_{k-1} \).

As a two-dimensional example, suppose that we have \( X = \{(0, 0), (1, 1)\} \subseteq \mathbb{Z} \). Then \( S(X, 2) = \{(0, 0), (1, 0), (0, 1), (1, 1), (2, 2), (2, 3), (3, 2), (3, 3)\} \), and we have \( \rho_2 : S(X, 2) \to X \) given by \( \rho(0, 0) = \rho(1, 0) = \rho(0, 1) = \rho(1, 1) = (0, 0) \), and \( \rho(2, 2) = \rho(2, 3) = \rho(3, 2) = \rho(3, 3) = (1, 1) \). Finally, in Fig. 3, we show the points of \( S(D, 2) \), with \( D \) the diamond as in Fig. 1 above, and indicate the way in which the points of \( S(D, 2) \) are aggregated by the projection \( \rho_2 : S(D, 2) \to D \).

**Fig. 2** Aggregation of points by \( \rho_3 : S(I_2, 3) = I_8 \to I_2 \)

**Fig. 3** Aggregation of points by \( \rho_2 : S(D, 2) \to D \)
Subdivision behaves well with respect to products. For any digital images $X \subseteq \mathbb{Z}^m$ and $Y \subseteq \mathbb{Z}^n$ and any $k \geq 2$ we have an isomorphism of digital images

$$S(X \times Y, k) \cong S(X, k) \times S(Y, k)$$

and, furthermore, the standard projection $\rho_k: S(X \times Y, k) \to X \times Y$ may be identified with the product of the standard projections on $X$ and $Y$, thus

$$\rho_k = \rho_k \times \rho_k: S(X, k) \times S(Y, k) \to X \times Y.$$  

Note also that we may iterate subdivision. It is straightforward to check that, for any $k, l \geq 1$, we have an isomorphism of digital images $S(S(X, k), l) \cong S(X, kl)$.

**Example 3.3** We mentioned above that, for $I_0 = \{0\} \subseteq \mathbb{Z}$, we have $S(I_0, k) = S(0, k) = I_{k-1}$. For the origin $0 = (0, \ldots, 0) \in \mathbb{Z}^n$, we have $S(0, k) = (I_{k-1})^n$, an $n$-cube in $\mathbb{Z}^n$, and we may identify the projection $\rho_k: (I_{k-1})^n \to \{0\}$ as a product of projections

$$\rho_k \times \cdots \times \rho_k: I_{k-1} \times \cdots \times I_{k-1} \to I_0 \times \cdots \times I_0.$$  

More generally, for any $x \in \mathbb{Z}$, we have

$$S(x, k) = [kx, kx + k - 1] = \{kx + r \mid 0 \leq r \leq k - 1\}.$$  

If $x = (x_1, \ldots, x_r) \in \mathbb{Z}^n$, then we have

$$S(x, k) = \{(kx_1 + r_1, \ldots, kx_n + r_n) \mid 0 \leq r_i \leq k - 1\} = [kx_1, kx_1 + k - 1] \times \cdots \times [kx_n, kx_n + k - 1] = S(x_1, k) \times \cdots \times S(x_n, k).$$  

These descriptions make plain that we may identify the projection $\rho_k: S(x, k) \to \{x\}$ with the product of projections

$$\rho_k \times \cdots \times \rho_k: S(x_1, k) \times \cdots \times S(x_n, k) \to \{x_1\} \times \cdots \times \{x_1\}.$$  

**Remarks 3.4** Various notions of subdivision of a digital image appear in the digital topology literature. The underlying idea in all versions is to enlarge a digital image, so as to allow more flexibility in maps and homotopies. But the way in which we approach this in Definition 3.1 differs somewhat from other versions. For example, [5,6] use the set we have denoted $S'(X, k)$ in Definition 3.1 (it is denoted by $S(X, k)$ in these papers), with a view towards using multivalued functions to achieve the desired flexibility. The approach there adapts the approach of [7,8], in which subdivisions of a form similar to what we would write as $S'(X, 3')$ are used, again with a view to using the multivalued functions introduced there. The extra step we take in Definition 3.1, to arrive at $S(X, k)$ from $S'(X, k)$, means that our subdivided digital image is again a digital image (points have integer coordinates), and we do not need to describe

\[ \mathbb{Z} \text{ Springer} \]
adjacency in $S(X, k)$ separately. Whilst our form of subdivision involves this extra step, it seems formally satisfactory to have the subdivision as a digital image and it allows us to avoid keeping track of which sub-multiples of integers are being used across different subdivisions.

By an inclusion of digital images (of the same dimension) $j: A \to X \subseteq \mathbb{Z}^n$ we mean that $A$ is a subset of $X$ (the coordinates of a point of $A$ remain the same under inclusion into $X$). It is easy to see that, given an inclusion of digital images $j: A \to X \subseteq \mathbb{Z}^n$, we have an obvious corresponding continuous inclusion of subdivisions $S(j, k): S(A, k) \to S(X, k)$ such that the diagram

$$
\begin{array}{ccc}
S(A, k) & \xrightarrow{S(j, k)} & S(X, k) \\
\rho_k \downarrow & & \rho_k \\
A & \xrightarrow{j} & X
\end{array}
$$

commutes. We say that the map $S(j, k)$ covers the map $j$. Indeed, we may give an explicit formula as follows. For each point $a \in A$, write $a = (a_1, \ldots, a_n)$. Also, write $t = (t_1, \ldots, t_n)$, with $0 \leq t_1, \ldots, t_n \leq k - 1$, for a typical point $t$ in the cubical $k \times k \times \cdots \times k$ lattice $(I_{k-1})^n \subseteq \mathbb{Z}^n$. Then the points of $S(a, k) \subseteq S(A, k)$ may be written as

$$
S(a, k) = \{k a + t \mid t \in (I_{k-1})^n\} = \{(k a_1 + t_1, \ldots, k a_n + t_n) \mid 0 \leq t_1, \ldots, t_n \leq k - 1\}
$$

with $\rho_k(k a + t) = a$ for all $t \in (I_{k-1})^n$. Here, the scalar multiple $k a$ and the sum $k a + t$ denote coordinate-wise (vector) scalar multiplication and addition in $\mathbb{Z}^n$. Then $S(j, k): S(A, k) \to S(X, k)$ may be written as

$$
S(j, k)(k a + t) = k j(a) + t,
$$

where $j(a) = (a_1, \ldots, a_n) \in X$. It is easy to confirm that this gives a (continuous) map.

For a more general map $f: X \to Y$, however, it is not so clear how we should construct a map of subdivisions that covers the map, in the sense of a filler—a map that occupies the place of the dotted arrow—for the following (commutative) diagram:

$$
\begin{array}{ccc}
S(X, k) & \xrightarrow{S(j, k)} & S(Y, k) \\
\rho_k \downarrow & & \rho_k \\
X & \xrightarrow{f} & Y
\end{array}
$$

In fact, it is not even obvious that such a map of subdivisions always exists, in general. In this paper we show that such a map does exist for $k$ odd and for arbitrary maps.
of digital images with 1D and 2D domains. However, as the next several examples illustrate, the formulation of (2) will not provide such a map in general.

**Example 3.5** Consider the constant map of 1D digital images \( c: I_1 \to I_0 = [0] \), given by \( c(1) = c(0) = 0 \). If we use the formulation of (2) above to define a function

\[
S(c, k): S(I_1, k) \to S(I_0, k)
\]

as \( S(c, k)(k a + t) = k c(a) + t \), then we have \( S(c, k)(k-1) = S(c, k)(k \cdot 0 + k - 1) = k c(0) + k - 1 = k - 1 \) but \( S(c, k)(k) = S(c, k)(k \cdot 1 + 0) = k c(1) + 0 = 0 \). Then \( k-1 \sim_{S(I_1, k)} k \) but \( k-1 \sim_{S(I_0, k)} 0 \) unless \( k = 2 \): the function \( S(c, k) \) is not continuous for \( k \geq 3 \). See Fig. 4 for an illustration of this situation.

In this example, defining \( C: S(I_1, k) \to S(I_0, k) \) as a constant map, \( C(k a + t) = 0 \), for instance, gives a continuous map that covers \( c \). But the point here is, that it is not obvious how to adapt a covering map of subdivisions depending on the given map.

The issue is not confined to functions that coalesce points together, either. Here are two examples of injective maps for which \( S(f, k) \), defined as in (2) above, fails to be continuous.

**Example 3.6** (a) Consider the map \( f: I_1 \to I_1 \) given by \( f(0) = 1 \) and \( f(1) = 0 \). The function \( S(f, k): S(I_1, k) = I_{2k-1} \to S(I - 1, k) = I_{2k-1} \) defined by the formulation of (2) above gives

\[
S(f, k)(k-1) = 2k - 1 \quad \text{and} \quad S(f, k)(k) = 0.
\]

For \( k \geq 2 \), we have \( 2k - 1 \sim 0 \) so this function is not continuous for any \( k \geq 2 \). See Fig. 5, in which we have indicated the way in which the projections \( \rho_k: S(I_1, k) \to I_1 \) aggregate points.

(b) (Similar to an observation illustrated in [6, Fig.1].) Consider the map \( f: X \to Y \) with \( X = \{(0, 0), (1, 0), (0, 1)\} \subseteq \mathbb{Z}^2 \), \( Y = \{(0, 0), (1, 1)\} \subseteq \mathbb{Z}^2 \), and \( f \) given by

\[
f(0, 0) = (0, 0), \quad f(1, 0) = (1, 1), \quad f(0, 1) = (1, 1).
\]

The function \( S(f, 2): S(X, 2) \to S(Y, 2) \) defined by the formulation of (2) above gives

\[
(0, 0) \mapsto (0, 0), \quad (1, 0) \mapsto (1, 0), \quad (0, 1) \mapsto (0, 1), \quad (1, 1) \mapsto (1, 1)
\]
on the four points of $S((0, 0), 2)$. Likewise for the four points in $S((1, 0), 2)$, $S(f, 2)$ would give

$$(2, 0) \leftrightarrow (2, 2), \quad (2, 1) \leftrightarrow (2, 3), \quad (3, 0) \leftrightarrow (3, 2), \quad (3, 1) \leftrightarrow (3, 3).$$

But this would result in adjacent points $(1, 0) \sim_{S(X, 2)} (2, 1)$ being mapped to non-adjacent points $(1, 0) \sim_{S(Y, 2)} (2, 3)$, for example. The situation is summarized in Fig. 6, in which we want a filler $S(X, 2) \to S(Y, 2)$ that makes the diagram commute. Notice one feature of this example, in particular. Although we have $f(0, 0) = (0, 0)$, it is not possible for a covering map of $f$ to restrict to the identity $S((0, 0), k) \to S((0, 0), k)$. For $k = 2$, for instance, we see in Fig. 6 that $(0, 1) \sim_{X} (0, 2)$ and $(0, 1) \sim_{X} (1, 2)$, but any covering map of $f$ must map both $(0, 2)$ and $(1, 2)$ to points of $S((1, 1), 2)$ in $S(Y, 2)$, none of which are adjacent to $(0, 1) \in S(Y, 2)$. That is, the possibilities for a covering map are constrained by how surrounding points are mapped by $f$, and not just by how the points themselves are mapped. In this example, it is not so clear how one should associate a continuous map $S(X, 2) \to S(Y, 2)$ to the original $f$, as part of a methodical scheme for doing so.

In the next two sections, we will give methodical constructions that, in particular, provide covering maps of subdivisions in the examples above. A more general question, special cases of which are also resolved in the following sections, is to ask how—or whether—a map of digital images of different dimensions might induce a covering map of subdivisions.

We close this section on subdivision with some constructions that we use in the following section and in the sequel. The projection $\rho_k: S(X, k) \to X$ may be factored—written as a composition—in various ways. For example, if $k = pq$, then we may write

$$\rho_k = \rho_p \circ \rho_q : S(X, k) \to S(X, p) \to X.$$
A different sort of “partial projection” that may also be used to factor $\rho_k$ is as follows.

**Definition 3.7** For any $x \in \mathbb{Z}$ and any $k \geq 2$, recall that the subdivision $S(x, k)$ may be described as $S(x, k) = [xk, xk + k - 1]$. Then, for $k \geq 3$, define a function

$$\rho_c^k : S(x, k) \to S(x, k - 1)$$

as

$$\rho_c^k(xk + j) = \begin{cases} xk + j & 0 \leq j \leq \lfloor k/2 \rfloor - 1, \\ xk + j - 1 & \lfloor k/2 \rfloor \leq j \leq k - 1. \end{cases}$$

Next, for any $x = (x_1, \ldots, x_n) \in \mathbb{Z}^n$, with the identifications from Example 3.3 of

$$S(x, k) = S(x_1, k) \times \cdots \times S(x_n, k) \quad \text{and} \quad S(x, k - 1) = S(x_1, k - 1) \times \cdots \times S(x_n, k - 1),$$

define $\rho_c^k : S(x, k) \to S(x, k - 1)$ as the product of functions

$$\rho_c^k \times \cdots \times \rho_c^k : S(x_1, k) \times \cdots \times S(x_n, k) \to S(x_1, k - 1) \times \cdots \times S(x_n, k - 1).$$
Finally, for any digital image $X \subseteq \mathbb{Z}^n$, define
\[ \rho^C_k : S(X, k) \to S(X, k - 1) \]
by viewing each subdivision as a (disjoint) union
\[ S(X, k) = \bigsqcup_{x \in X} S(x, k) \quad \text{and} \quad S(X, k - 1) = \bigsqcup_{x \in X} S(x, k - 1) \]
and assembling a global $\rho^C_k$ on $S(X, k)$ from the individual $\rho^C_k : S(x, k) \to S(x, k - 1)$ as just defined.

**Proposition 3.8** For $k \geq 3$, the partial projection $\rho^C_k : S(X, k) \to S(X, k - 1)$ is continuous.

**Proof** For $x \in \mathbb{Z}$, the map of intervals $\rho^C_k : S(x, k) \to S(x, k - 1)$ is easily seen to be continuous. Then, for any $x \in \mathbb{Z}^n$, we have defined $\rho^C_k : S(x, k) \to S(x, k - 1)$ as the product of individually continuous functions, hence it is also continuous. It remains to confirm that the $\rho^C_k$ assemble together to give a globally continuous function on $S(X, k)$.

So suppose that we have $y \in S(x, k)$ and $y' \in S(x', k)$ with $y \sim_{S(X,k)} y'$ and $x \neq x' \in X$. Note, though, that we must have $x \sim_X x'$, since $\rho^C_k(y) = x$, $\rho^C_k(y') = x'$, and $\rho^C_k : S(X, k) \to X$ is continuous. Write $x = (x_1, \ldots, x_n)$ and $x' = (x'_1, \ldots, x'_n)$. Then we have
\[ y = (kx_1 + r_1, \ldots, kx_n + r_n) \quad \text{and} \quad y' = (kx'_1 + r'_1, \ldots, kx'_n + r'_n) \]
for $r_i, r'_i \in \mathbb{Z}$ with $0 \leq r_i, r'_i \leq k - 1$, each $i = 1, \ldots, n$. Now for $\rho^C_k(y) \sim_{S(X,k-1)} \rho^C_k(y')$, it is necessary and sufficient that we have $\rho^C_k(kx_i + r_i) \sim \rho^C_k(kx'_i + r'_i)$ in $S(x_i, k - 1) \sqcup S(x'_i, k - 1) \subseteq \mathbb{Z}$, for each $i$. Write $\rho^C_k(kx_i + r_i) = (k - 1)x_i + s_i$ and $\rho^C_k(kx'_i + r'_i) = (k - 1)x'_i + s'_i$, with the $s_i, s'_i$ satisfying $0 \leq s_i, s'_i \leq k - 2$ and determined as in Definition 3.7. Then
\[ \rho^C_k(kx_i + r_i) - \rho^C_k(kx'_i + r'_i) = (k - 1)(x_i - x'_i) + (s_i - s'_i), \]
and we must show that, for each $i = 1, \ldots, n$, we have
\[ -1 \leq (k - 1)(x_i - x'_i) + (s_i - s'_i) \leq 1. \tag{3} \]

Because we have $x \sim_X x'$, it follows that, for each $i = 1, \ldots, n$, we have $-1 \leq x_i - x'_i \leq 1$. For each $i$, there are three possibilities. First, suppose that we have $x_i - x'_i = 1$. Then $y \sim_{S(X,k)} y'$ entails that, in the $i$th coordinates, we have
\[ 1 \geq kx_i + r_i - (kx'_i + r'_i) = k(x_i - x'_i) + (r_i - r'_i) = k + r_i - r'_i. \]
Thus $r'_i \geq k - 1 + r_i$ and the only possibility is that, in this coordinate, we have $r_i = 0$ and $r'_i = k - 1$. From Definition 3.7, then, we have $s_i = 0$ and $s'_i = k - 2$ and hence

\[ 0 \leq (k - 1)(x_i - x'_i) + (s_i - s'_i) \leq 1. \]
(k − 1)(x_i − x_i′) + (s_i − s_i′) = k − 1 − (k − 2) = 1, which satisfies (3). Second, suppose that we have x_i − x_i′ = −1. Then

\[-1 \leq kx_i + ri − (kx_i′ + r_i′) = k(x_i − x_i′) + (ri − r_i′) = −k + ri − r_i′,\]

thus r_i ≥ k − 1 + r_i′, and we have r_i′ = 0 and r_i = k − 1. From Definition 3.7, then, we have s_i′ = 0 and s_i = k − 2 and in this case (k − 1)(x_i − x_i′) + (s_i − s_i′) = −(k − 1) + (k − 2) = −1, which also satisfies (3). Finally, suppose that we have x_i − x_i′ = 0. Here, y ∼_{S(X,k)} y′ entails that we have

\[-1 \leq kx_i + ri − (kx_i′ + r_i′) = k(x_i − x_i′) + (ri − r_i′) = 0 + ri − r_i′ \leq 1,\]

so that r_i and r_i′ differ by at most 1. From Definition 3.7, if \{r_i, r_i′\} ⊆ [0, ⌊k/2⌋ − 1] or if \{r_i, r_i′\} ⊆ ([k/2], k − 1], then we have s_i − s_i′ = r_i − r_i′ and so −1 ≤ s_i − s_i′ ≤ 1. The only other possibility is that we have \{r_i, r_i′\} = ([k/2], ⌈k/2⌉] in which case s_i = s_i′ and so s_i − s_i′ = 0. Wherever r_i and r_i′ fall in [0, k − 1], then, we have (k − 1)(x_i − x_i′) + (s_i − s_i′) = s_i − s_i′ which satisfies |s_i − s_i′| ≤ 1 and (3) is again satisfied. The result follows. □

**Corollary 3.9**  Let \(X \subseteq \mathbb{Z}^n\) be any digital image. For any \(k \geq 3\), we may factor the projection \(\rho_k : S(X,k) → X\) as

\[\rho_k = \rho_{k−1} \circ \rho_k^c : S(X,k) → S(X,k−1) → X\]

with \(\rho_{k−1} : S(X,k−1) → X\) the standard projection and \(\rho_k^c : S(X,k) → S(X,k−1)\) the partial projection map from Definition 3.7.

**Proof**  It suffices to check that the composition agrees with \(\rho_k\) on \(S(x,k) \subseteq S(X,k)\), for each \(x \in X\). But when restricted to \(S(x,k)\), both \(\rho_k\) and \(\rho_{k−1} \circ \rho_k^c\) are constant maps. □

## 4 One-Dimensional Domains: Paths in \(Y\)

For \(Y \subseteq \mathbb{Z}^n\) a digital image and any \(N \geq 1\), a **path of length \(N\) in \(Y\)** is a continuous map \(α : I_N → Y\). Unlike in the ordinary (topological) homotopy setting, where any path may be taken with the fixed domain \([0, 1]\), in the digital setting we must allow paths to have different domains. Recall from Example 3.2 that we obtain a longer interval when we subdivide an interval: \(S(I_N,k) = I_{N(k+1)−k+1} \subseteq \mathbb{Z}\). In the following result, notice that the map of subdivisions that covers the given path is itself a path (of length \(N(2k + 1) + 2k\)) in the subdivided digital image \(S(Y,2k + 1)\).

**Theorem 4.1**  Suppose we are given \(α : I_N → Y\), a path of length \(N\) in any digital image \(Y \subseteq \mathbb{Z}^n\). For any integer \(k \geq 1\) there is a canonical choice of map of (odd) subdivisions

\[\widetilde{α} : S(I_N,2k+1) = I_{N(2k+1)+2k} → S(Y,2k+1)\]
that covers the given path, in the sense that the following diagram commutes:

\[
\begin{array}{c}
S(I_N, 2k + 1) \xrightarrow{\alpha} S(Y, 2k + 1) \\
\downarrow \rho_{2k+1} \hspace{5cm} \downarrow \rho_{2k+1} \\
I_N \xrightarrow{\hat{\alpha}} \hat{Y}.
\end{array}
\]

Our proof consists of a direct construction of the path \(\hat{\alpha}\). We first establish some notation and vocabulary used in the proof. If \(i \in I_N \subseteq \mathbb{Z}\), write

\[\overline{t} = (2k + 1)i + k \in S(i, 2k + 1).\]

So \(\overline{t}\) is the point in the center of the length \(2k\) subinterval \(S(i, 2k + 1) \subseteq S(I_N, 2k + 1)\) and, in particular, we have \(\rho_{2k+1}(\overline{t}) = i\), with \(\rho_{2k+1}: S(I_N, 2k + 1) \rightarrow I_N\) the standard projection. Then the \(2k + 1\) points of each \(S(i, 2k + 1)\) may be described as

\[S(i, 2k + 1) = \{\overline{t} + r \mid -k \leq r \leq k\}.\]

To describe points of \(S(Y, 2k + 1)\), we use notation similar to that used around (2) above in the discussion of covering an inclusion. Write \(y \in Y \subseteq \mathbb{Z}^n\) as \(y = (y_1, \ldots, y_n)\). Then, write

\[\overline{y} = ((2k + 1)y_1 + k, \ldots, (2k + 1)y_n + k) \in S(y, 2k + 1)\]

so that \(\overline{y}\) is the point in the center of \(S(y, 2k + 1)\), which is a cubical \((2k + 1) \times (2k + 1) \times \cdots \times (2k + 1)\) lattice in \(\mathbb{Z}^n\), namely the translate of \((I_{2k})^n \subseteq \mathbb{Z}^n\) by \((2k + 1)y\). Here, the scalar multiple \((2k + 1)y\) means coordinate-wise (vector) scalar multiplication; we will use coordinate-wise (vector) scalar multiplication and addition in \(\mathbb{Z}^n\) freely in our notation. Note, in particular, that we have \(\rho_{2k+1}(\overline{y}) = y\), with \(\rho_{2k+1}: S(Y, 2k + 1) \rightarrow Y\) the standard projection. Then the \((2k + 1)^n\) points of each \(S(y, 2k + 1)\) may be described as

\[S(y, 2k + 1) = \{\overline{y} + (r_1, \ldots, r_n) \mid -k \leq r_1, \ldots, r_n \leq k\}.\]

**Proof of Theorem 4.1** We will define our map of subdivisions

\[\hat{\alpha}: S(I_N, 2k + 1) \rightarrow S(Y, 2k + 1)\]

so that we have

\[\hat{\alpha}(i) = \alpha(i)\]

for each \(i \in I_N\). That is, for each \(i\), we will map the center of the subinterval \(S(i, 2k+1) \subseteq S(I_N, 2k+1)\) to the center of the cubical sub-lattice \(S(\alpha(i), 2k + 1) \subseteq S(Y, 2k + 1)\). For each \(i = 0, \ldots, N - 1\), we write

\[\alpha(i + 1) - \alpha(i)\]
for the displacement vector in \( \mathbb{Z}^n \) from \( \alpha(i) \) to \( \alpha(i + 1) \) in \( Y \). Recall that we intend coordinate-wise (vector) difference here. Since we have \( \alpha(i) \sim_{\gamma} \alpha(i + 1) \), each coordinate of \( \alpha(i + 1) - \alpha(i) \) is 0, 1, or -1. Also, let \( k = (k, \ldots, k) \in \mathbb{Z}^n \) denote the vector each of whose coordinates is \( k \). Then, for each \( i = 0, \ldots, N - 1 \), we have

\[
\alpha(i + 1) - \alpha(i) = (2k + 1)\alpha(i + 1) + k - ((2k + 1)\alpha(i) + k) = (2k + 1)[\alpha(i + 1) - \alpha(i)].
\]

(8)

Write \( S(I_N, 2k + 1) = I_{(2k+1)N+2k} \) as a union of subintervals

\[
[0, (2k + 1)N + 2k] = [0, \bar{0}] \cup \left\{ \bigcup_{i=0}^{N-1} [\overline{i}, \overline{i+1}] \right\} \cup [\overline{N}, (2k + 1)N + 2k],
\]

(9)

where we have \( \overline{i} = i(2k + 1) + k \), whence \( \overline{i+1} = \overline{i} + 2k + 1 \). Then for \( j \in S(I_N, 2k + 1) \) define

\[
\hat{\alpha}(j) := \begin{cases} 
\alpha(0) & 0 \leq j \leq \bar{0}, \\
\frac{\alpha(i) + (j - \overline{i})[\alpha(i + 1) - \alpha(i)]}{\alpha(N)} & \overline{i} \leq j \leq \overline{i+1} \text{ with } i \in I_N, \\
\alpha(i) & \bar{N} \leq j \leq (2k + 1)N + 2k.
\end{cases}
\]

(10)

If \( j = \bar{0} = [0, \bar{0}] \cap [\overline{0}, \overline{1}] \), then this formula clearly gives a well-defined value of \( \alpha(0) \). For \( j = \overline{i} \) with \( i = 1, \ldots, N \), we have \( j = [\overline{i-1}, \overline{i}] \cap [\overline{i}, \overline{i+1}] \) if \( i \neq N \) and \( j = [\overline{N-1}, \overline{N}] \cap [\overline{N}, (2k + 1)N + 2k] \) if \( i = N \), and the formula also gives a well-defined value of \( \alpha(i) \) as follows from (8) (with \( i \) replaced by \( i - 1 \) in that identity). Otherwise, \( j \) falls into a unique subinterval of the decomposition (9) and it follows that \( \hat{\alpha} \) is well-defined on \( S(I_N, 2k + 1) \) and also maps centers to centers as in (7). Indeed, all we have really done in this definition is to map the \( 2k + 2 \) points of \( [\overline{i}, \overline{i+1}] \subseteq S(I_N, 2k + 1) \) so as to interpolate the segment between the centers of the lattices \( S(\alpha(i), 2k + 1) \) and \( S(\alpha(i + 1), 2k + 1) \) in \( S(Y, 2k + 1) \), for each \( i \).

To check continuity of \( \hat{\alpha} \) it is sufficient to check continuity when restricted to each of the subintervals of the decomposition (9) separately. This is because consecutive subintervals in (9) overlap in their endpoints, and thus any two consecutive integers in \( S(I_N, 2k + 1) \) both lie in a single subinterval. On \([0, \bar{0}]\) and \([\overline{N}, (2k + 1)N + 2k]\), \( \hat{\alpha} \) is defined to be constant and so is obviously continuous when restricted to either of these subintervals. So suppose we have consecutive integers \( \{j, j + 1\} \subseteq [\overline{i}, \overline{i+1}] \) for some \( i \in I_N \). The definition of \( \hat{\alpha} \) in (10) gives

\[
\hat{\alpha}(j + 1) - \hat{\alpha}(j) = \alpha(i) + ((j + 1) - i)[\alpha(i + 1) - \alpha(i)] - (\alpha(i) + (j - i)[\alpha(i + 1) - \alpha(i)]) = \alpha(i + 1) - \alpha(i).
\]

Now the fact that each \( \alpha(i + 1) - \alpha(i) \) has coordinates taken from \{0, \pm1\} implies that \( \hat{\alpha}(j + 1) \) and \( \hat{\alpha}(j) \) are adjacent in \( S(Y, 2k + 1) \), and continuity of \( \hat{\alpha} \) follows.
Finally, for diagram (4) to commute, it is necessary and sufficient that the map \( \hat{\alpha} \) satisfy

\[
\hat{\alpha}(S(i, 2k + 1)) \subseteq S(\alpha(i), 2k + 1) \subseteq S(Y, 2k + 1)
\]

for each \( i \in I_N \). To confirm this, divide each \( S(i, 2k + 1) \), for \( i \in I_N \), into a (disjoint) union of subintervals

\[
S(i, 2k + 1) = [(2k + 1)i, (2k + 1)i + k - 1] \cup [\bar{t}, \bar{t} + k]
\]

with the first subinterval consisting of the \( k \) points to the left of the center \( \bar{t} \) and the second consisting of the \( k + 1 \) points to the right (including the center \( \bar{t} \) itself). If \( 1 \leq i \leq N \), then we may re-write the decomposition (12) as

\[
S(i, 2k + 1) = [\bar{i} - 1 + k + 1, \bar{i} - 1 + 2k] \cup [\bar{t}, \bar{t} + k].
\]

First consider \( j \in [\bar{t}, \bar{t} + k] \), a point in the right-hand subinterval of (12) or (13) for any \( i \in I_N \). Then (10) gives

\[
\hat{\alpha}(j) = \alpha(\bar{i}) + (j - \bar{t})[\alpha(i + 1) - \alpha(i)].
\]

Since \( j \in [\bar{t}, \bar{t} + k] \), we have \( 0 \leq j - \bar{t} \leq k \). It follows from (6) (with \( y \) replaced by \( \alpha(i) \)) that \( \hat{\alpha}([\bar{t}, \bar{t} + k]) \subseteq S(\alpha(i), 2k + 1) \) because \( \alpha(i + 1) - \alpha(i) \) has coordinates taken from \( \{0, \pm 1\} \).

It remains to check (11) holds for points in the left-hand subinterval of (12) or (13). First take \( i = 0 \) in (12). Then (10) gives \( \hat{\alpha}([0, k - 1]) = \alpha(0) \) so we have \( \hat{\alpha}([0, k - 1]) \subseteq S(\alpha(0), 2k + 1) \). Now take \( 1 \leq i \leq N \), but with the left-hand subinterval written as in (13). Then for \( j \in [\bar{i} - 1 + k + 1, \bar{i} - 1 + 2k] \) with \( i \in [1, N] \), (10) gives

\[
\hat{\alpha}(j) = \alpha(i - 1) + (j - \bar{i} - 1)[\alpha(i) - \alpha(i - 1)]
\]

with \( k + 1 \leq (j - \bar{i} - 1) \leq 2k \). But we may re-write \( \bar{i} - 1 \) as

\[
\bar{i} - 1 = \bar{t} - (2k + 1)
\]

and from (8) (with \( i \) replaced by \( i - 1 \)) we may re-write \( \alpha(i - 1) \) as

\[
\alpha(i - 1) = \alpha(i) - (2k + 1)[\alpha(i) - \alpha(i - 1)].
\]

Substituting the last two identities into (14) yields

\[
\hat{\alpha}(j) = \alpha(i) + (j - \bar{t})[\alpha(i) - \alpha(i - 1)]
\]

with the constraint on \( j \) written \( -k \leq j - \bar{t} \leq -1 \). As in the argument for the right-hand subinterval of (12), it now follows from (6) that we have

\[
\hat{\alpha}([\bar{i} - 1 + k + 1, \bar{i} - 1 + 2k]) \subseteq S(\alpha(i), 2k + 1).\]
Combining the last few steps shows that (11) holds for each $i \in I_N$. □

**Example 4.2** Return to part (a) of Example 3.6 and consider the map $f : I_1 \to I_1$ given by $f(0) = 1$ and $f(1) = 0$. Applying Theorem 4.1, we obtain a map

$$\widehat{\alpha} : S(I_1, 3) = I_5 \to S(I_1, 3) = I_5$$

that covers $f$. It is given by $\widehat{\alpha}(i) = 5 - i$ for $1 \leq i \leq 4$, and $\widehat{\alpha}(0) = 4$, $\widehat{\alpha}(5) = 1$.

**Example 4.3** Let $\alpha : I_N \to Y$ be a constant path in $Y \subseteq \mathbb{Z}^n$. Suppose that we have $\alpha(i) = y_0 \in Y$ for $0 \leq i \leq N$. For any odd $2k + 1$, the map $\widehat{\alpha} : S(I_N, 2k + 1) \to S(Y, 2k + 1)$ given by Theorem 4.1 that covers $\alpha$ is simply the constant path $\widehat{\alpha}(j) = \overline{y_0} \in S(Y, 2k + 1)$ for $0 \leq j \leq (2k + 1)N + 2k$. For instance, we would cover the constant map in Example 3.5 with the constant map $\widehat{\alpha} : S(I_1, 3) = I_5 \to S(I_0, 3) = I_2$ with $\widehat{\alpha}(i) = 1 \in I_2$ for each $i \in I_5$.

We will refer to the cover $\widehat{\alpha}$ of a path $\alpha$ constructed in Theorem 4.1 as the **standard cover** of the path. Ideally, we would like to construct a functorial cover of maps of digital images regardless of the dimension of the domain, but we are not able to do so at present. We observe here, though, that the standard cover of a path does have some functorial-like properties, such as the following:

**Lemma 4.4** Let $Y \subseteq \mathbb{Z}^n$ be any digital image. For any path $\alpha : I_N \to Y$, let $\widehat{\alpha}$ denote the standard cover with respect to $(2k + 1)$-fold subdivisions, so that $\widehat{\alpha}$ makes the following diagram commute:

$$
\begin{array}{ccc}
S(I_N, 2k + 1) & \xrightarrow{\widehat{\alpha}} & S(Y, 2k + 1) \\
\rho_{2k+1} \downarrow & & \downarrow \rho_{2k+1} \\
I_N & \xrightarrow{\alpha} & Y.
\end{array}
$$

(a) If $C_N : I_N \to Y$ denotes the constant path at a point $y_0 \in Y$, then we have $\widehat{C}_N = C_N' : I_N' \to S(Y, 2k + 1)$, the constant path at $\overline{y_0}$, where $N' = (2k + 1)N + 2k$.

(b) If $Y = I_N$ and $\alpha : I_N \to I_N$ is the identity, then we have

$$
\widehat{id}_{I_N} = id_{S(I_N, 2k + 1)} : S(I_N, 2k + 1) \to S(I_N, 2k + 1).
$$

**Proof** Both parts follow from a careful reading of the definition of $\widehat{\alpha}$. □

## 5 Two-Dimensional Domains: Surfaces in $Y$

Our goal in this section is to generalize the results of the previous section and describe, for a given map $H : I_M \times I_N \to Y$ and a choice of $k \geq 1$, a map $\widehat{H}$ for which the following diagram commutes:
At the end of this section, we generalize to maps \( f : X \to Y \) with \( X \) a general 2D digital image and not just a rectangle.

Before embarking on our construction of the map \( \widehat{H} \), we assemble a few ingredients that we will use. We continue to use the notation established before Theorem 4.1 and also during its proof. Suppose we have a map

\[
H : I_M \times I_N \to Y
\]

for integers \( M, N \geq 1 \). For each \( t \) with \( 0 \leq t \leq N \), define \( \alpha_t : I_M \to Y \), and for each \( s \) with \( 0 \leq s \leq M \), define \( \beta_s : I_N \to Y \) as

\[
\alpha_t(s) = H(s, t) \text{ for } 0 \leq s \leq M \quad \text{and} \quad \beta_s(t) = H(s, t) \text{ for } 0 \leq t \leq N.
\]

So the \( \alpha_t \) are the horizontal, and the \( \beta_s \) are the vertical, coordinate curves of \( H \). Choose a positive integer \( k \). Then as in Theorem 4.1, each of these paths has a standard cover

\[
\widehat{\alpha}_t : S(I_M, 2k + 1) = I_{(2k+1)M+2k} \to S(Y, 2k + 1) \quad \text{and} \quad \widehat{\beta}_s : S(I_N, 2k + 1) = I_{(2k+1)N+2k} \to S(Y, 2k + 1).
\]

We will eventually define \( \widehat{H} \) in such a way as to have these be amongst the horizontal and vertical coordinate curves of \( \widehat{H} \), respectively.

Recall from our generalities on subdivision in Sect. 3 that we have an isomorphism of digital images \( S(I_M \times I_N, 2k + 1) \cong S(I_M, 2k + 1) \times S(I_N, 2k + 1) \). For individual points \( (i, j) \in I_M \times I_N \), we may specialize this identification to an isomorphism \( S((i, j), 2k + 1) \cong S(i, 2k + 1) \times S(j, 2k + 1) \). We use these identifications repeatedly in what follows.

Recall also from Theorem 4.1 that, for \( i \in I_M \), we write the center of the subinterval \( S(i, 2k + 1) \subseteq S(I_M, 2k + 1) \) as \( \overline{i} = i(2k + 1) + k \). Then each \( (2k + 1) \times (2k + 1) \) sub-lattice \( S((i, j), 2k + 1) \subseteq S(I_M \times I_N, 2k + 1) \) has the point

\[
\overline{(i, j)} := (\overline{i}, \overline{j}) = (i(2k + 1) + k, j(2k + 1) + k)
\]

at its center. We refer to these points as \textit{centers} of the sub-divided digital image \( S(I_M \times I_N, 2k + 1) \). Furthermore, for a point \( y \in Y \), we write \( \overline{y} \) for the center of \( S(y, 2k + 1) \subseteq S(Y, 2k + 1) \).

For \( a, b \in \mathbb{Z} \) with \( a < b \), let \([a, b] := \{ x \in \mathbb{Z} \mid a \leq x \leq b \}\) denote a general interval (as opposed to \( I_N \)). Then \([a, b] \times [c, d], \) with \( a < b \) and \( c < d \) denotes a rectangle in \( \mathbb{Z}^2 \) (or square if \( b - a = d - c \)). We write \( \partial ([a, b] \times [c, d]) \) for the \textit{boundary of the rectangle} \([a, b] \times [c, d] \), by which we mean the points

\[\square\text{ Springer}\]
\[ \partial([a, b] \times [c, d]) := \{(x, y) \in [a, b] \times [c, d] \mid x \in [a, b] \text{ or } y \in [c, d]\}. \]

For a given \( k \geq 1 \), the \( n \)-cube \([\overline{0}, \overline{1}]^n = [k, 3k+1]^n\) has a central \( 2^n \)-clique. Namely, the \( 2^n \)-clique at the center of the larger cube that consists of the unit \( n \)-cube \([\overline{0}+k, \overline{0}+k+1]^n = [2k, 2k+1]^n\). For instance, the central 4-clique of \([\overline{0}, \overline{1}]^2\) consists of the four points

\[ \{(2k, 2k), (2k+1, 2k), (2k, 2k+1), (2k+1, 2k+1)\}. \]

Also, we may divide the \( n \)-cube \([\overline{0}, \overline{1}]^n\) into \( 2^n \) sub-cubes, or orthants (quadrant if \( n = 2 \)) as we will refer to them in the sequel, consisting of products of \( n \) intervals

\[ J_1 \times \cdots \times J_n \]

with each interval \( J_r \) equal to \([\overline{0}, \overline{2k}] = [k, 2k] \text{ or } [2k+1, \overline{1}] = [2k+1, 3k+1] \). Each orthant contains exactly one corner of the \( n \)-cube \([\overline{0}, \overline{1}]^n\)—\((c_1, \ldots, c_n)\), say, with each coordinate \( c_r \) equal to \( k \) or \( 3k+1 \)—and exactly one point of the central clique—\((\gamma_1, \ldots, \gamma_n)\), say, with each coordinate \( \gamma_r \) equal to \( 2k \) or \( 2k+1 \). The coordinates of these two points are related, in that \((c_r, \gamma_r)\) equals either \((k, 2k)\) or \((3k+1, 2k+1)\) for each \( r = 1, \ldots, n \). These features of the cube \([\overline{0}, \overline{1}]^n\) may be translated to any \( n \)-cube of side-length \( 2k+1 \). Suppose \( a_1, \ldots, a_n \) are integers and we use them to form the \( n \)-cube

\[ \overline{A} = [\overline{a_1}, \overline{a_1+1}] \times [\overline{a_2}, \overline{a_2+1}] \times \cdots \times [\overline{a_n}, \overline{a_n+1}] \]
\[ = [a_1(2k+1)+k, a_1(2k+1)+3k+1] \times \cdots \]
\[ \times [a_n(2k+1)+k, a_n(2k+1)+3k+1]. \]

Then \( \overline{A} \) is the translate of \([\overline{0}, \overline{1}]^n\) by the translation vector \((2k+1)(a_1, \ldots, a_n)\); the orthants and central clique of \([\overline{0}, \overline{1}]^n\) translate to those of \( \overline{A} \). Specifically, the central clique of \( \overline{A} \) consists of the unit \( n \)-cube

\[ [a_1(2k+1)+2k, a_1(2k+1)+2k+1] \times \cdots \]
\[ \times [a_n(2k+1)+2k, a_n(2k+1)+2k+1] \]

and the orthants of \( \overline{A} \) consist of \( n \)-cubes \( J_1 \times \cdots \times J_n \) with each interval \( J_r \) equal to \([\overline{a_r}, \overline{a_r+k}] = [a_r(2k+1)+k, a_r(2k+1)+2k] \text{ or } [\overline{a_r+k}, \overline{a_r+k+2k+1}] = [a_r(2k+1)+2k+1, a_r(2k+1)+3k+1]\).

At various points in what follows, it is convenient to break a function into its coordinate functions and work with them individually. We mean this as follows. Suppose given a map \( F: X \to Y \) with \( Y \) a product \( Y = Y_1 \times \cdots \times Y_n \). Then for each \( x \in X \), we may write

\[ F(x) = (F_1(x), \ldots, F_n(x)). \]
where $F_i : X \to Y_i$ is the $i$th coordinate function of $F$. It is immediate from the definitions that $F$ is continuous if and only if each of its coordinate functions is continuous.

We will use the following construction at several points in the developments of this section.

**Definition 5.1** Let $Y \subseteq \mathbb{Z}^n$ be any digital image and let $i, j$ be integers. Given a map $H : [i, i + 1] \times [j, j + 1] \to Y$ from a unit square to $Y$, define a unit $n$-cube $A_H \subseteq \mathbb{Z}^n$ as follows. On each of the four points of $[i, i + 1] \times [j, j + 1]$ write $H$ coordinate-wise as

$$H(i, j) = (H_1(i, j), \ldots, H_n(i, j)),$$

and so on for the other three points. Then, for each coordinate $r = 1, \ldots, n$, set

$$a_r = \min \{H_r(i, j), H_r(i + 1, j), H_r(i, j + 1), H_r(i + 1, j + 1)\},$$

and let $A_H$ denote the unit $n$-cube

$$A_H := [a_1, a_1 + 1] \times \cdots \times [a_n, a_n + 1] \subseteq \mathbb{Z}^n.$$

Now for a given integer $k \geq 1$ let $\overline{A_H}$ denote the $n$-cube

$$\overline{A_H} = [\overline{a_1}, \overline{a_1} + 1] \times [\overline{a_2}, \overline{a_2} + 1] \times \cdots \times [\overline{a_n}, \overline{a_n} + 1]$$

so that $\overline{A_H}$ is an $n$-cube with side of length $2k + 1$. Because $a_r$ is the minimum of $\{H_r(i, j), H_r(i + 1, j), H_r(i, j + 1), H_r(i + 1, j + 1)\}$, and because $H$ is continuous, it follows that we have

$$a_r \leq H_r(i, j), H_r(i + 1, j), H_r(i, j + 1), H_r(i + 1, j + 1) \leq a_r + 1$$

for each $r$ and for each point of the 4-clique $[i, i + 1] \times [j, j + 1]$. So each of the four points $\{H(i, j), H(i + 1, j), H(i, j + 1), H(i + 1, j + 1)\} \subseteq Y$ are amongst the corners of $A_H$, but some corners of $A_H$ may lie outside $Y$. Likewise, each of the four points $\{H(i, j), H(i + 1, j), H(i, j + 1), H(i + 1, j + 1)\} \subseteq S(Y, 2k + 1)$ are amongst the corners of $\overline{A_H}$, but some corners of $\overline{A_H}$ may lie outside $S(Y, 2k + 1)$. Notice that the standard projection restricts to a map $\rho_{2k+1} : \overline{A_H} \to A_H$ under which each orthant of $\overline{A_H}$ is mapped to a corner of $A_H$.

Now we start to construct the map $\hat{H}$ of (16). Define a subset $L_{M \times N} \subseteq S(I_M \times I_N, 2k + 1)$, the lattice of horizontals and verticals through the centers of the rectangle, as

$$L_{M \times N} := \{(s, \overline{j}) \mid s \in S(I_M, 2k + 1), \; j \in I_N\} \cup \{\overline{i}, t) \mid i \in I_M, \; t \in S(I_N, 2k + 1)\}.$$
We define a map \( \hat{H} : L_M \times N \rightarrow S(Y, 2k+1) \) by setting, for each \( i \in I_M \) and \( j \in I_N \),

\[
\hat{H}(i, t) := \hat{\beta}_i(t) \quad \text{for } 0 \leq t \leq (2k+1)N + 2k,
\]

\[
\hat{H}(s, j) := \hat{\alpha}_j(s) \quad \text{for } 0 \leq s \leq (2k+1)M + 2k.
\]

In Fig. 7, dots represent points on which we have now defined \( \hat{H} \). Notice the dashed lines that divide each square \([i, i+1] \times [j, j+1]\) into four quadrants. (These lines are not integer gridlines. Rather, they pass between points with integer coordinates.) Each such quadrant contains a center of \( S(I_M \times I_N, 2k+1) \) (the solid dot) and all points in one of these quadrants of \([i, i+1] \times [j, j+1]\) are mapped to one point of the 4-clique \([i, i+1] \times [j, j+1] \subseteq I_M \times I_N \) by the standard projection.

**Lemma 5.2** Given \( H : I_M \times I_N \rightarrow Y \) with \( Y \subseteq \mathbb{Z}^n \) and a choice of \( k \geq 1 \), define \( \hat{H} : L_M \times N \rightarrow S(Y, 2k+1) \) as in (17) above. Then \( \hat{H} \) is continuous and the following diagram commutes:

\[
\begin{array}{ccc}
L_M \times N & \xrightarrow{\hat{H}} & S(Y, 2k+1) \\
\rho_{2k+1} \downarrow & & \downarrow \rho_{2k+1} \\
I_M \times I_N & \xrightarrow{H} & Y,
\end{array}
\]

where the left-hand \( \rho_{2k+1} \) denotes the restriction of \( \rho_{2k+1} : S(I_M \times I_N, 2k+1) \rightarrow I_M \times I_N \).

**Proof** Note that \( \hat{H} \) is well defined at any point in the lattice \( L_M \times N \) at which a horizontal row or vertical column intersect. This is because, for any \( (i, j) \in I_M \times N \), we have

\[
\{ (s, j) \mid s \in S(I_M, 2k+1) \} \cap \{ (i, t) \mid t \in S(I_N, 2k+1) \} = (i, j),
\]

which is a center. Now the way in which we defined the standard cover of a path in Theorem 4.1 extended the definition of (7), so we have

\[
\hat{\alpha}_j(i) = \alpha_j(i) = H(i, j) \quad \text{and} \quad \hat{\beta}_i(j) = \beta_i(j) = H(i, j).
\]
Since these agree, \( \hat{H} \) takes the well-defined value \( \overline{H}(i, j) \) at the intersection point \((i, j) \in L_{M \times N}\). There is no issue with \( \hat{H} \) being well defined at any other points of \( L_{M \times N} \).

When \( \hat{H} \) is restricted to a horizontal row of \( L_{M \times N} \), we have

\[
\rho_{2k+1} \circ \hat{H} = H \circ \rho_{2k+1} : \{(s, j) \mid s \in S(I_M, 2k+1)\} \to Y,
\]

for each \( j \in I_N \). This follows since we defined \( \hat{H}(s, j) = \hat{a}_j(s) \) for \( s \in S(I_M, 2k+1) \) with \( \hat{a}_j(s) = H(s, j) \) for \( s \in I_M \), and \( \rho_{2k+1} \circ \hat{a}_j = \alpha_j \circ \rho_{2k+1} : S(I_M, 2k+1) \to Y \).

Similarly, when \( \hat{H} \) is restricted to a vertical column of \( L_{M \times N} \), we have

\[
\rho_{2k+1} \circ \hat{H} = H \circ \rho_{2k+1} : \{(i, t) \mid t \in S(I_N, 2k+1)\} \to Y,
\]

for each \( i \in I_M \). Since each point of \( L_{M \times N} \) lies in a horizontal row or a vertical column, it follows that diagram (18) commutes.

Now we check continuity of \( \hat{H} \). When restricted to a horizontal row or to a vertical column of \( L_{M \times N} \), continuity of \( \hat{H} \) follows immediately from the continuity of the standard covers \( \hat{a}_j \) and \( \hat{b}_i \). It remains to check that \( \hat{H}(s, t) \sim \hat{H}(s', t') \) for adjacent points \((s, t) \sim (s', t') \in L_{M \times N}\) with \((s, t)\) on a horizontal and \((s', t')\) on a vertical. Namely, for points

\[
(s, t) = (i \pm 1, j) \quad \text{and} \quad (s', t') = (i, j \pm 1),
\]

for each \( i \in I_M \) and \( j \in I_N \). We separate such pairs of points into one of two types:

(a) \( s \in \{0 - 1, \overline{M} + 1\} \) or \( t' \in \{0 - 1, \overline{N} + 1\} \);
(b) \( s \in \{0 + 1, \overline{M} - 1 \mid 1 \leq i \leq M - 1\} \) and \( t' \in \{0 + 1, \overline{N} + 1 - 1 \mid 1 \leq j \leq N - 1\} \).

We argue separately for each type of pair. For type (5), suppose we have \( (s, t) = (0 - 1, j) \) and \( (s', t') = (0, j \pm 1) \). Then from the definition of \( \hat{a}_j \), we have \( \hat{H}(0 - 1, j) = \hat{a}_j(0 - 1) = \alpha_j(0) = \hat{H}(0, j) \).

But \( \hat{H}(0, j) = \hat{b}_0(j \pm 1) = \hat{H}(0, j \pm 1) \), from the continuity of \( \hat{b}_0 \). It follows that we have \( \hat{H}(s, t) \sim \hat{H}(s', t') \) here. Similar arguments confirm that we have \( \hat{H}(s, t) \sim \hat{H}(s', t') \) when \((s, t) \sim (s', t')\) and \( s = \overline{M} + 1 \) or \( t' \in \{0 - 1, \overline{N} + 1\} \).

For pairs \((s, t) \sim (s', t')\) of type (5), note that any such pair of points lie on two edges of a square with corners \((i, j), \overline{i} + 1, \overline{j}, j \) for some \((i, j) \) with \( 0 \leq i \leq M - 1 \) and \( 0 \leq j \leq N - 1 \). Such a square is illustrated in Fig. 8.

Here, dots indicate points of \( L_{M \times N} \), on which we have defined \( \hat{H} \). Notice the dashed lines preserved from Fig. 7. The adjacent points of type (5) in this square are circled.

So given \((i, j)\) with \( 0 \leq i \leq M - 1 \) and \( 0 \leq j \leq N - 1 \), use the given \( H \) restricted to \([i, i + 1] \times [j, j + 1]\) to determine a unit \( n \)-cube \( A_H \) and an \( n \)-cube \( \overline{A_H} \) with side of length \( 2k + 1 \), both in \( \mathbb{Z}^n \), as in Definition 5.1. Some points of the \( n \)-cube \( \overline{A_H} \) may lie outside \( S(Y, 2k + 1) \), but because diagram (18) commutes, the image under \( \hat{H} \) of \( \partial([i, i + 1] \times [j, j + 1]) \), the boundary of the square \([i, i + 1] \times [j, j + 1]\), does lie in \( S(Y, 2k + 1) \). Furthermore, we have defined \( \hat{H} \) in such a way that corners of \( \partial([i, i + 1] \times [j, j + 1]) \) are mapped to corners of \( \overline{A_H} \) and the edges of \( \partial([i, i + 1] \times [j, j + 1]) \).
Fig. 8 Four pairs of adjacent points of type (b) in a square $[i, i + 1] \times [j, j + 1]$. Illustrated with $k = 5$ or $2k + 1 = 11$

$[j, j + 1]$ are mapped to edges or diagonals of $\overline{A_H}$ that join corner to corner. Now in a cube, such as $\overline{A_H}$, any two points adjacent to a corner of the cube must be adjacent to each other, as the corner point and all its neighbors form a $2^n$-clique. It follows that for a pair $(s, t) \sim (s', t')$ of type (b), we have $\tilde{H}(s, t) \sim \tilde{H}(s', t')$ since both of these image points lie in the cube $A_H$ and each is adjacent to a corner of $\overline{A_H}$. We have checked that $\tilde{H}$ preserves adjacency for any pair of adjacent points in $L_M \times N$. Thus $\tilde{H}$ is a continuous map.

Now we wish to extend the map $\tilde{H}: L_M \times N \to Y$ over the whole rectangle $I_M \times I_N$ to obtain a map $\tilde{H}: S(I_M \times I_N, 2k + 1) \to Y$ such that diagram (16) commutes. We begin by establishing a partial special case, which will form a key part of our argument for the general result. Take $M = N = 1$, so that $I_M = I_N = I_1 = [0, 1]$, take $Y = (I_1)^n \subseteq \mathbb{Z}^n$, and suppose given a map $H: I_1 \times I_1 \to (I_1)^n$ and a choice of $k \geq 1$. Rather than construct a map $\tilde{H}: S(I_1 \times I_1, 2k + 1) \to S((I_1)^n, 2k + 1)$ here, we focus on the sub-square $[0, 1] \times [0, 1] \subseteq S(I_1 \times I_1, 2k + 1)$ and construct a map from this sub-square into the sub-cube $[0, 1]^n$ of $S((I_1)^n, 2k + 1)$. The reasons for this focus will eventually emerge when we prove our result for general rectangular domains.

Our extension of a map from the boundary to the whole square makes use of the following device.

**Definition 5.3 (coordinate-centering function)** For each $k \geq 1$, consider the subinterval $[\overline{0}, \overline{1}] = [k, 3k + 1] \subseteq S(I_1, 2k + 1)$. Define the map $C_k: [\overline{0}, \overline{1}] \to [\overline{0}, \overline{1}]$ by

$$
C_k(x) = \begin{cases} 
  x + 1 & \text{if } \overline{0} = k \leq x \leq 2k - 1 = \overline{0} + k - 1, \\
  x & \text{if } x = \overline{0} + k = 2k \text{ or } x = \overline{0} + k + 1 = 2k + 1, \\
  x - 1 & \text{if } \overline{0} + k + 2 = 2k + 2 \leq x \leq 3k + 1 = \overline{1}.
\end{cases}
$$
We refer to this map $C_k$ as the coordinate-centering function (for a given $k$). Although $C_k$ depends on the choice of $k$, we often use the notation $C$ for $C_k$ in situations in which the choice of $k$ is understood.

The coordinate-centering function plays a prominent role in much that follows. The idea is that $C$ may be used to progressively move each coordinate of a point of an $n$-cube closer to that of a point in the central $2^n$-clique of the cube, as described above Definition 5.1 at the start of this section. For instance, in $[\widetilde{0}, \widetilde{1}]^n$ (for a given $k$) define a function $c: \{k, 3k+1\} \to \{2k, 2k+1\}$ by $c(k) = 2k$ (or $c(\widetilde{0}) = \widetilde{0} + k$) and $c(3k + 1) = 2k + 1$ (or $c(\widetilde{1}) = \widetilde{0} + k + 1$). Then for each corner of $[\widetilde{0}, \widetilde{1}]^n$,

$$(y_1, \ldots, y_n) \text{ with } \{y_1, \ldots, y_n\} \subseteq \{\widetilde{0}, \widetilde{1}\},$$

the closest point to that corner in the central clique of $[\widetilde{0}, \widetilde{1}]^n$ is $(c(y_1), \ldots, c(y_n))$. By iterating the coordinate-centering function, we may obtain the same result: for each coordinate of the corner point, we have $C^k(y_i) = c(y_i)$, where the notation $C^k$ denotes the $k$th power or iterate of $C$. Indeed, we can parametrize the path in $[\widetilde{0}, \widetilde{1}]^n$ from corner to closest central-clique point as

$$\{(C^s(y_1), \ldots, C^s(y_n)) \mid s = 0, \ldots, k\}, \quad (19)$$

where again the notation $C^s$ denotes the $s$th power or iterate of $C$ and we mean $C^0(y_i) = y_i$. Then the points (19) constitute a diagonal of an orthant from (outside) corner to opposite (central) corner of an orthant $[\widetilde{0}, \widetilde{1}]^n$.

In the next result we coordinatize points in the square $[\widetilde{0}, \widetilde{1}]^2$ in terms of a point on the boundary of the square together with the number of steps needed to arrive at the point from the boundary, using the function defined in Definition 5.3.

**Lemma 5.4** For an integer $k \geq 1$, quarter the square $[k, 3k + 1]^2$ into quadrants $J_1 \times J_2$ with each $J_1, J_2$ equal to $[k, 2k]$ or $[2k + 1, 3k + 1]$. On each quadrant, define a function

$$D: J_1 \times J_2 \to \partial([k, 3k + 1]^2) \times [0, k]$$

as follows:

- On $[k, 2k]^2$: with $r := \min \{R \mid (s - R, t - R) \in \partial([k, 3k + 1]^2)\}$, set $D(s, t) := (s - r, t - r, r)$.
- On $[k, 2k] \times [2k + 1, 3k + 1]$: with $r := \min \{R \mid (s - R, t + R) \in \partial([k, 3k + 1]^2)\}$, set $D(s, t) := (s - r, t + r, r)$.
- On $[2k + 1, 3k + 1] \times [k, 2k]$: with $r := \min \{R \mid (s + R, t - R) \in \partial([k, 3k + 1]^2)\}$, set $D(s, t) := (s + r, t - r, r)$.
- On $[2k + 1, 3k + 1]^2$: with $r := \min \{R \mid (s + R, t + R) \in \partial([k, 3k + 1]^2)\}$, set $D(s, t) := (s + r, t + r, r)$.

(Note that $D$ is not a continuous function on the square $[k, 3k + 1]^2$.) Then, for each $(s, t) \in [k, 3k + 1]^2$, if $D(s, t) = (s_b, t_b, r)$, then we have

$$(s, t) = (C^r(s_b), C^r(t_b)).$$
Fig. 9 Indicating how points interior to a square $[t, t+1] \times [j, j+1]$ are associated to a boundary point, together with the number of steps from the boundary to that point. Illustrated with $k = 5$ or $2k + 1 = 11$.

**Proof** Points $(s, t)$ on the boundary $\partial([k, 3k+1]^2)$ have $D(s, t) = (s, t, 0)$. For points not on the boundary, work quadrant-by-quadrant. For example, consider a point $(s, t) \in [k+1, 2k]^2$, namely in the quadrant $[k, 2k]^2$ but not on the boundary. Any such point satisfies $t = s + m$ for some $m$ with $1 - k \leq m \leq k - 1$. Then there is a unique point $(s_b, t_b) \in [k, 2k]^2 \cap \partial([k, 3k+1]^2)$ whose coordinates also satisfy $t = s + m$. Then

\[
\{(s_b, t_b), (C(s_b), C(t_b)), (C^2(s_b), C^2(t_b)), \ldots, (C^r(s_b), C^r(t_b))\}
\]  

is a sequence of adjacent points in $[k, 2k]^2$ whose coordinates satisfy $t = s + m$, starting at $(s_b, t_b)$ and ending at $(s, t) = (C^r(s_b), C^r(t_b))$. This follows because, in the quadrant $[k, 2k]^2$, the effect of $C$ is to increase each coordinate by 1 until one or other coordinate equals $2k$, and the choice of $r$ that we made for $(s, t)$ is such that we have $C^{r-1}(s_b) < 2k$ and $C^{r-1}(t_b) < 2k$. The same reasoning, mutatis mutandis, may be applied in the other three quadrants. In Fig. 9 we have indicated the paths that contain segments of type (20) in each quadrant. Dots indicate points in the boundary $\partial([k, 3k+1]^2)$; solid dots indicate centers. Numbers indicate points not on the boundary. For such a point, if $D(s, t) = (s_b, t_b, r)$, then the point is labeled with an $r$, and connected to its boundary point $(s_b, t_b)$ by the segment that contains $(s, t)$. Notice that for $(s, t) \in [k+1, 3k]^2$, namely not on the boundary $\partial([k, 3k+1]^2)$, if $D(s, t) = (s_b, t_b, r)$, then $s_b \notin \{2k, 2k+1\}$ and $t_b \notin \{2k, 2k+1\}$. 

For the next result, recall that $I_1$ denotes the unit interval $I_1 = [0, 1]$, consisting of two points.

**Lemma 5.5** Let $H : (I_1)^2 \to (I_1)^n$ be any map. Choose an integer $k \geq 1$. On the boundary $\partial([0, 1]^2)$ of the square $[0, 1]^2 = [k, 3k+1]^2 \subseteq S((I_1)^2, 2k+1)$, define a map

\[ F : \partial([0, 1]^2 \times [0, 1]^2) \to [0, 1]^n \subseteq \mathbb{Z}^n \]

by restricting the map $\hat{H} : L_{1 \times 1} \to S((I_1)^n, 2k+1)$ of Lemma 5.2. Namely, with reference to the definitions in (17) and (10), define
\[
F(\bar{0}, t) := \overline{H(0, 0)} + t[H(0, 1) - H(0, 0)], \\
F(\bar{1}, t) := \overline{H(1, 0)} + t[H(1, 1) - H(1, 0)], \\
F(s, \bar{0}) := \overline{H(0, 0)} + s[H(1, 0) - H(0, 0)], \\
F(s, \bar{1}) := \overline{H(0, 1)} + s[H(1, 1) - H(0, 1)], \\
\]

for \(0 \leq s, t \leq 2k + 1\). Then extend \(F\) over \([\bar{0}, \bar{1}]^2\) as follows: For \((s, t) \in [\bar{0}, \bar{1}]^2\), let \(D(s, t) = (s_b, t_b, r)\) as in Lemma 5.4, where \((s_b, t_b) \in \partial([\bar{0}, \bar{1}]^2)\) and \(0 \leq r \leq k\) with \((s, t) = (C^r(s_b), C^r(t_b))\). Define \(F(s, t)\) coordinate-wise as

\[
F_i(s, t) := C^r(F_i(s_b, t_b))
\]

for each coordinate \(i = 1, \ldots, n\). Then \(F\) is a continuous map \(F : [\bar{0}, \bar{1}]^2 \to [\bar{0}, \bar{1}]^n\) that restricts to \(F\) on the boundary and makes the following diagram commute:

\[
\begin{array}{ccc}
[\bar{0}, \bar{1}] \times [\bar{0}, \bar{1}] & \xrightarrow{F} & [\bar{0}, \bar{1}]^n \\
\rho_{2k+1} \downarrow & & \downarrow \rho_{2k+1} \\
[0, 1] \times [0, 1] & \xrightarrow{H} & [0, 1]^n.
\end{array}
\]

**Proof** It is automatic that the diagram (23) commutes. This is because the original \(F\) on the boundary is a map that covers \(H\) (see diagram (18)—\(F\) on the boundary is a restriction of \(\overline{H}\)), and (22) entails that, on any quadrant of \([\bar{0}, \bar{1}]^2\), the images under \(F\) of points in that quadrant lie in the same orthant of \([\bar{0}, \bar{1}]^n\) as do the images of the points on the boundary in that orthant.

To establish continuity of the extended \(F\), we work by induction on \(k\). Induction starts with \(k = 1\), where \([\bar{0}, \bar{1}] = [1, 4]\) and we are given \(F\) on \(\partial([1, 4]^2)\). Extending over \([1, 4]^2\) entails assigning values to \(F\) on \([2, 3]^2 \subseteq [1, 4]^2\). With reference to Lemma 5.4, we have

\[
D(2, 2) = (1, 1, 1), \quad D(2, 3) = (1, 4, 1), \\
D(3, 2) = (4, 1, 1), \quad \text{and} \quad D(3, 3) = (4, 4, 1).
\]

So (22) gives

\[
F(2, 2) := (C(F_1(1, 1)), C(F_2(1, 1)), \ldots, C(F_n(1, 1))), \\
F(2, 3) := (C(F_1(1, 4)), C(F_2(1, 4)), \ldots, C(F_n(1, 4))), \\
F(3, 2) := (C(F_1(4, 1)), C(F_2(4, 1)), \ldots, C(F_n(4, 1))), \\
F(3, 3) := (C(F_1(4, 4)), C(F_2(4, 4)), \ldots, C(F_n(4, 4))).
\]

Continuity requires that we have \(F(s, t) \sim F(s', t')\) in \([1, 4]^n\) whenever \((s, t) \sim (s', t')\) in \([1, 4]^2\). If both \((s, t)\) and \((s', t')\) lie in \(\partial([1, 4]^2)\), then continuity of the original given \(F\) guarantees the result. If both \((s, t)\) and \((s', t')\) lie in \([2, 3]^2\), then

\[\square\] Springer
the comments following Definition 5.3 entail that \( F(s, t) \) and \( F(s', t') \) both lie in the central \( 2^n \)-clique of \([1, 4]^n\), namely \([2, 3]^n\) (this is also clear from the formula for the extended \( F \)), and thus are adjacent. So it only remains to check for \( (s, t) \in [2, 3]^2 \) and \((s', t') \in \partial([1, 4]^2)\). There are four such points \((s, t)\), namely \((2, 2), (2, 3), (3, 2), \) and \((3, 3)\). The argument for each point is the same, \( \textit{mutatis mutandis} \), so we just argue with \((s, t) = (2, 2)\). For this choice of \((s, t)\), if \((s', t') \in \{(1, 1), (1, 2), (2, 1)\}\), then both \( F(s, t) \) and \( F(s', t') \) lie in the same orthant of \([1, 4]^n\) as each other. But every orthant of \([1, 4]^n\) is a \( 2^n \)-clique, so \( F(s, t) \sim F(s', t') \) here. The only other points \((s', t')\) adjacent to \((s, t) = (2, 2)\) are \((s', t') = (1, 3)\) and \((s', t') = (3, 1)\). Again the argument for each choice is the same \( \textit{mutatis mutandis} \), so we just argue for \((s', t') = (1, 3)\). From their definitions, we have that \( F(1, 1) \) is some corner of \([1, 4]^n\) and \( F(2, 2) \) is the unique member of the central clique \([2, 3]^n \subseteq [1, 4]^n\) that is in the orthant determined by the corner \( F(1, 1) \). Said differently, \( F(2, 2) \) is the unique member of the central clique of \([1, 4]^n\) that is adjacent to the corner \( F(1, 1) \). Then the points of \([1, 4]^n\) adjacent to \( F(2, 2) \) form an \( n \)-cube of side-length 2 (three points on each side) that has \( F(1, 1) \) as a corner. Since \( F \) is defined along the side of \( \partial([1, 4]^2) \) from \((1, 1)\) to \((1, 4)\) as a path (an edge or diagonal of \([1, 4]^n\) in fact), there must be a path in \([1, 4]^n\) of length 2 that starts at \( F(1, 1) \) and ends at \( F(1, 3) \). But all such paths are contained the cube of points adjacent to \( F(2, 2) \), and so we have \( F(1, 3) \sim F(2, 2) \). This is sufficient to conclude that \( F \) is continuous for \( k = 1 \).

Now assume inductively that continuity holds for some \( k \geq 1 \), and suppose we have \( F \) defined on \( \partial([0, 1]^2) \) as in the enunciation but for \( k + 1 \), so that now \([0, 1] = [k + 1, 3k + 4] \) and \([0, 1]^2 \) is a square of side-length \( 2k + 3 \). Our plan is to check continuity on \( \partial([k + 1, 3k + 4]^2) \cup \partial([k + 2, 3k + 3]^2) \), and then use the induction hypothesis to conclude continuity on \([k + 2, 3k + 3]^2 \). But in order to do this, we need to match the definition of \( F \) from (22) with the form from Lemma 5.2 on the smaller square \([k + 2, 3k + 3]^2 \). To this end, we write out details for the left-hand edges of the squares. The other edges co-operate similarly, and we omit details for them. So first note that on the lower-left corner of \( \partial([k + 1, 3k + 4]^2) \), since \( F \) here is the restriction of \( \widehat{H} \), we have

\[
F_i(k + 1, k + 1) = \overline{H_i(0, 0)} = \begin{cases} 
  k + 1 & \text{if } H_i(0, 0) = 0, \\
  3k + 4 & \text{if } H_i(0, 0) = 1.
\end{cases}
\]

Hence on the lower-left corner of \( \partial([k + 2, 3k + 3]^2) \), (22) gives

\[
F_i(k + 2, k + 2) = \begin{cases} 
  C(k + 1) = k + 2 & \text{if } H_i(0, 0) = 0, \\
  C(3k + 4) = 3k + 3 & \text{if } H_i(0, 0) = 1.
\end{cases}
\]  \hfill (24)

Then up the left-hand edge of \( \partial([k + 2, 3k + 3]^2) \), (22) gives

\[
F_i(k + 2, k + 2 + t) = \begin{cases} 
  C(F_i(k + 1, k + 1 + t)) & 0 \leq t \leq k, \\
  C(F_i(k + 1, k + 3 + t)) & k + 1 \leq t \leq 2k + 1.
\end{cases}
\]

\( \Box \) Springer
Again, since $F$ on $\partial([k + 1, 3k + 4]^2)$ is the restriction of $\hat{H}$, we may re-write the term $C(F_i(k + 1, k + 1 + t))$ for $0 \leq t \leq k$ as

$$C(F_i(k + 1, k + 1 + t)) = C \left( \overline{H_i(0, 0)} + t[H_i(0, 1) - H_i(0, 0)] \right)$$

$$= \begin{cases} 
C(k + 1) = k + 2 & H_i(0, 0) = 0 = H_i(0, 1), \\
C(k + 1 + t) = k + 2 + t & H_i(0, 0) = 0, \; H_i(0, 1) = 1, \\
C(3k + 4 - t) = 3k + 3 - t & H_i(0, 0) = 1, \; H_i(0, 1) = 0, \\
C(3k + 4) = 3k + 3 & H_i(0, 0) = 1 = H_i(0, 1).
\end{cases} \quad (25)$$

Notice that here $C$ denotes the coordinate-centering function $C_k : [k + 1, 3k + 4] \to [k + 1, 3k + 4]$, and so increases values by 1 in the range $[k + 1, 2k + 1]$ and decreases by 1 in the range $[2k + 4, 3k + 4]$. For $0 \leq t \leq k$, we have $k + 1 + t \in [k + 1, 2k + 1]$ and $3k + 4 - t \in [2k + 4, 3k + 4]$, so the last set of terms are correct.

Similarly, we may re-write $C(F_i(k + 1, k + 3 + t))$ for $k + 1 \leq t \leq 2k + 1$ as

$$C(F_i(k + 1, k + 3 + (t + 2))) = C \left( \overline{H_i(0, 0)} + (t + 2)[H_i(0, 1) - H_i(0, 0)] \right)$$

$$= \begin{cases} 
C(k + 1) = k + 2 & H_i(0, 0) = 0 = H_i(0, 1), \\
C(k + 1 + (t + 2)) = k + 2 + t & H_i(0, 0) = 0, \; H_i(0, 1) = 1, \\
C(3k + 4 - (t + 2)) = 3k + 3 - t & H_i(0, 0) = 1, \; H_i(0, 1) = 0, \\
C(3k + 4) = 3k + 3 & H_i(0, 0) = 1 = H_i(0, 1).
\end{cases} \quad (26)$$

The middle two terms here follow as, with $k + 1 \leq t \leq 2k + 1$, we have $k + 1 + (t + 2) = k + 3 + t \in [2k + 4, 3k + 4]$ and $3k + 4 - (t + 2) = 3k + 2 - t \in [k + 1, 2k + 1]$. But now notice that the final forms of (25) and (26) agree, so with (24) we may write, for $F$ on the left-hand edge of $\partial([k + 2, 3k + 3]^2)$,

$$F(k + 2, k + 2 + t) = F(k + 2, k + 2) + t[H(0, 1) - H(0, 0)].$$

Similar reasoning on the other three edges of $[k + 2, 3k + 3]^2$ (we omit the details) shows that we may re-write $F$ on $\partial([k + 2, 3k + 3]^2)$ as

$$F(k + 2, k + 2 + t) = F(k + 2, k + 2) + t[H(0, 1) - H(0, 0)],$$

$$F(k + 2 + s, k + 2) = F(k + 2, k + 2) + s[H(1, 0) - H(0, 0)],$$

$$F(3k + 3, k + 2 + t) = F(3k + 3, k + 2) + t[H(1, 1) - H(1, 0)],$$

$$F(k + 2 + s, 3k + 3) = F(k + 2, 3k + 3) + s[H(1, 1) - H(0, 1)],$$

for $0 \leq s \leq 2k + 1$ and $0 \leq t \leq 2k + 1$.

So when restricted to the smaller square, $F$ does indeed have the aspect of a standard cover of $H$, as given in the equations (21). We make the identification precise as follows. Let $T : \mathbb{Z} \to \mathbb{Z}$ be the translation given by $T(x) = x + 2$ for $x \in \mathbb{Z}$. Then $T : [k, 3k + 1] \to [k + 2, 3k + 3]$. The reverse translation $T^{-1}(x) = x - 2$ for $x \in \mathbb{Z}$
is the inverse of \(T\), and we may use it to define a map

\[
G := (T^{-1})^n \circ F \circ T^2 : [k, 3k + 1]^2 \to [k, 3k + 1]^n,
\]

where \(F\) is defined in (22). Continuing to give details for the left-hand edge, we may write, for \((k, k + t) \in \partial ([k, 3k + 1]^2),\)

\[
G_i(k, k + t) = T^{-1} \circ F_i(k + 2, k + 2 + t)
= T^{-1}(F_i(k + 2, k + 2 + t) + t[H_i(0, 1) - H_i(0, 0)])
= T^{-1}(F_i(k + 2, k + 2)) + t[H_i(0, 1) - H_i(0, 0)]
\]

for \(0 \leq t \leq 2k + 1\). Now from (24) we have

\[
(H_i(0, 0))_k = T^{-1}(F_i(k + 2, k + 2)) = \begin{cases} k & H_i(0, 0) = 0, \\ 3k + 1 & H_i(0, 0) = 1, \end{cases}
\]

where the notation \((H_i(0, 0))_k\) emphasizes that here we mean the center with respect to \(k\): we have \((0)\)_k = k and \((1)\)_k = 3k + 1. It follows that we have

\[
G(k, k + 1) = (H(0, 0))_k + t[H(0, 1) - H(0, 0)].
\]

Similar reasoning (which again, we omit) along the other three sides shows that \(G\) agrees with the restriction of \(\hat{H}\) (with respect to \(k\)) on the whole of \(\partial ([k, 3k + 1]^2)\). To analyze \(G\) over the whole of \([k, 3k + 1]^2\), we use the function \(D\) from Lemma 5.4 with the subscripts \(D_k\) and \(D_{k+1}\) to indicate whether we are operating on \([k, 3k + 1]^2\) or on \([k + 1, 3k + 3]^2\). If \((s, t) \in [k, 3k + 1]^2\), then in the notation of Lemma 5.4 and the present result, we have \(D_{k+1}(s + 2, t + 2) = (s_b, t_b, r)\) for a unique \((s_b, t_b) \in \partial ([k + 1, 3k + 4]^2)\) and \(1 \leq r \leq k + 1\). Then \((C(s_b), C(t_b)) \in \partial ([k + 2, 3k + 3]^2)\) and satisfies \((C^{-1}(C(s_b) - 2), C^{-1}(C(t_b))) = (s + 2, t + 2)\). Then we have

\[
D_k(s, t) = (C(s_b) - 2, C(t_b) - 2, r - 1).
\]

Our definition of \(G\) (and the definition of \(F\)) gives

\[
G_i(s, t) = T^{-1} \circ F_i \circ T^2(s, t) = T^{-1} \circ F_i(s + 2, t + t)
= T^{-1}\circ C_{k+1}^{-1}(F_i(s_b, t_b)) = T^{-1}\circ C_{k+1}^{-1}C_{k+1}(F_i(s_b, t_b))
= T^{-1}\circ C_{k+1}^{-1}(F_i(C(s_b), C(t_b))) = C_{k+1}^{-1}T^{-1}(F_i(C(s_b), C(t_b))),
\]

where we have written \(C_{k+1} : [k + 1, 3k + 4] \to [k + 1, 3k + 4]\) for the coordinate-centering function with respect to \(k + 1\) and \(C_k : [k, 3k + 1] \to [k, 3k + 1]\) for the same with respect to \(k\). Note the identity \(T^{-1} \circ C_{k+1} = C_k \circ T^{-1} : [k + 1, 3k + 4] \to [k, 3k + 1]\) used in the last step. On the other hand, for \((C(s_b) - 2, C(t_b) - 2) \in \)
\[ \partial([k, 3k + 1]^2), \text{we have} \]
\[ G_i(C(s_b) - 2, C(t_b) - 2) = T^{-1}(F_i(C(s_b), C(t_b))). \]

From the last two displayed formulas, we have
\[ G_i(s, t) = C_k^{-1}(G_i(C(s_b) - 2, C(t_b) - 2)) \]
for a point \((s, t) \in [k, 3k + 1]^2\) with data as in Lemma 5.4 given by \(D_k(s, t) = (C(s_b) - 2, C(t_b) - 2, r - 1)\). Thus we may apply our induction hypothesis to \(G\) and conclude it is a continuous map. But then we have
\[ F = T^n \circ G \circ (T^{-1})^2 : [k + 2, 3k + 3]^2 \rightarrow [k + 2, 3k + 3]^n, \]
which is continuous since it is a composition of continuous maps.

To complete our inductive step we argue that \(F\) restricts to a continuous map
\[ F : \partial([k + 1, 3k + 4]^2) \sqcup \partial([k + 2, 3k + 3]^2) \rightarrow [k + 1, 3k + 4]^n. \]

Since our inductive hypothesis gives that \(F\) is continuous on \(\partial([k + 1, 3k + 4]^2)\), and we just argued that \(F\) is continuous on \([k + 2, 3k + 3]^2\), it remains to show that \(F\) preserves adjacency for points, one of which is in \(\partial([k + 1, 3k + 4]^2)\) and the other of which is in \(\partial([k + 2, 3k + 3]^2)\). Consider such a pair on the left-hand edge first. Suppose we have \((k + 1, k + 1 + t') \in \partial([k + 1, 3k + 4]^2)\) and \((k + 2, k + 2 + t) \in \partial([k + 2, 3k + 3]^2)\). Then \(0 \leq t \leq k + 1\) and adjacency of these points implies that we have \(t' \in \{t, t + 1, t + 2\}\), or \(1 + t - t' \in \{0, \pm 1\}\). We may express \(F(k + 1, t')\) coordinate-wise, in the style of (27), as
\[ F_i(k + 1, k + 1 + t') = \begin{cases} 
  k + 1 & H_i(0, 0) = 0 = H_i(0, 1), \\
  k + 1 + t' & H_i(0, 0) = 0 \text{ and } H_i(0, 1) = 1, \\
  3k + 4 - t' & H_i(0, 0) = 1 \text{ and } H_i(0, 1) = 0, \\
  3k + 4 & H_i(0, 0) = 1 = H_i(0, 1). 
\end{cases} \] (28)

for \(0 \leq t' \leq 2k + 3\). On the other hand, from the discussion leading up to (27) (see (25) and (26)), we have
\[ F_i(k + 2, k + 2 + t) = \begin{cases} 
  k + 2 & H_i(0, 0) = 0 = H_i(0, 1), \\
  k + 2 + t & H_i(0, 0) = 0 \text{ and } H_i(0, 1) = 1, \\
  3k + 3 - t & H_i(0, 0) = 1 \text{ and } H_i(0, 1) = 0, \\
  3k + 3 & H_i(0, 0) = 1 = H_i(0, 1). 
\end{cases} \] (29)

Comparing (28) and (29) in each of their four cases shows that we have \(F_i(k + 1, k + 1 + t') - F_i(k + 2, k + 2 + t) \in \{0, \pm 1\}\) for each \(i\), and thus \(F(k + 1, k + 1 + t') \sim F(k + 2, k + 2 + t)\). Arguing similarly for the top, bottom and right edges it is easy...
to confirm that $F$ preserves adjacency amongst pairs of points, one of which is in $\partial([k + 1, 3 k + 4]^2)$ and the other of which is in $\partial([k + 2, 3 k + 3]^2)$. However, any such pair of points is contained in one of these edges, and it follows that $F$ is continuous on the (disjoint) union $\partial([k + 1, 3 k + 4]^2) \sqcup \partial([k + 2, 3 k + 3]^2)$.

We have now confirmed $F$ is continuous on $\partial([k + 1, 3 k + 4]^2) \sqcup \partial([k + 2, 3 k + 3]^2)$ and on $[k + 2, 3 k + 3]^2$ separately. But this is sufficient to conclude that $F$ is continuous as a map $[k + 1, 3 k + 4]^2$, since any pair of adjacent points in $[k + 1, 3 k + 4]^2$ is contained either in $[k + 2, 3 k + 3]^2$ or in $\partial([k + 1, 3 k + 4]^2) \sqcup \partial([k + 2, 3 k + 3]^2)$. This completes the inductive step; the result follows. □

**Corollary 5.6** For $i$ and $j$ integers, let $[i, i + 1] \times [j, j + 1]$ be any unit square in $\mathbb{Z}^2$. Suppose we have a map $g: [i, i + 1] \times [j, j + 1] \to Y$, with $Y \subseteq \mathbb{Z}^n$ any digital image. Choose an integer $k \geq 1$ and define a map

$$G: \partial([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]) \to S(Y, 2k + 1)$$

by the formulas

$$G(\bar{i}, \bar{j} + t) := \overline{g(i, j) + t[g(i, j + 1) - g(i, j)]},$$
$$G(\bar{i} + 1, \bar{j} + t) := \overline{g(i + 1, j) + t[g(i + 1, j + 1) - g(i + 1, j)]},$$
$$G(\bar{i} + s, \bar{j}) := \overline{g(i, j) + s[g(i + 1, j) - g(i, j)]},$$
$$G(\bar{i} + s, \bar{j} + 1) := \overline{g(i, j + 1) + s[g(i + 1, j + 1) - g(i, j + 1)]}$$

for $0 \leq s, t \leq 2k + 1$. Then there is a canonical choice of an extension of $G$ over the whole square such that the following diagram commutes:

$$\begin{array}{ccc}
[i, i + 1] \times [j, j + 1] & \xrightarrow{\rho_{2k+1}} & S(Y, 2k + 1) \\
\downarrow G & & \downarrow \rho_{2k+1} \\
[i, i + 1] \times [j, j + 1] & \xrightarrow{g} & Y.
\end{array}$$

Here, $\rho_{2k+1}: [\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1] \to [i, i + 1] \times [j, j + 1]$ denotes the restriction of the standard projection $\rho_{2k+1}: S([i, i + 1] \times [j, j + 1], 2k + 1) \to [i, i + 1] \times [j, j + 1]$, either to the square $[\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]$ or to its boundary. Furthermore, the map $G$ is constructed in such a way that the image of the central 4-clique of $[\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]$ under $G$ is contained in the central $2^n$-clique of the $n$-cube $\overline{A_g}$ determined by $g$ as in Definition 5.1.

**Proof** We will apply Lemma 5.5 to extend $G$ over the interior of this square. To do so, use $g$ to determine a unit $n$-cube

$$A_g := [a_1, a_1 + 1] \times \cdots \times [a_n, a_n + 1] \subseteq \mathbb{Z}^n$$
as in Definition 5.1. Then write
\[ A_g := [\overline{a_1}, \overline{a_1} + 1] \times [\overline{a_2}, \overline{a_2} + 1] \times \cdots \times [\overline{a_n}, \overline{a_n} + 1] \subseteq \mathbb{Z}^n. \]

We will assemble a few ingredients that combine into diagram (32) below. Define translations \(T_1, \overline{T}_1 : \mathbb{Z}^2 \to \mathbb{Z}^2\) by
\[
T_1(x_1, x_2) = (x_1 - i, x_2 - j) \quad \text{and} \quad \overline{T}_1(x_1, x_2) = (x_1 - (2k + 1)i, x_2 - (2k + 1)j)
\]
so that, in particular, we have \(T_1(i, j) = (0, 0)\) and \(\overline{T}_1(i, j) = (\overline{0}, \overline{0})\). Likewise, define translations \(T_2, \overline{T}_2 : \mathbb{Z}^n \to \mathbb{Z}^n\) by
\[
T_2(x_1, \ldots, x_n) = (x_1 - a_1, \ldots, x_n - a_n) \quad \text{and} \quad \overline{T}_2(x_1, \ldots, x_n) = (x_1 - (2k + 1)a_1, \ldots, x_n - (2k + 1)a_n)
\]
so that we have \(T_2(a_1, \ldots, a_n) = (0, 0, \ldots, 0)\) and \(\overline{T}_2(\overline{a_1}, \ldots, \overline{a_n}) = (\overline{0}, \overline{0}, \ldots, \overline{0})\). We encountered \(\overline{T}_2\) (or its inverse) in the discussion above Definition 5.1: it translates \(A_g\) to \([\overline{0}, 1]^n\). As may be seen in diagram (32), these translations together translate the given data into the situation of Lemma 5.5. Translations \(\overline{T}_1\) and \(\overline{T}_2\) preserve the boundaries of the squares or \(n\)-cubes; both pairs of translations respect standard projections, in that we have
\[
\rho_{2k+1} \circ \overline{T}_1 = T_1 \circ \rho_{2k+1} : [\overline{i}, i + 1] \times [\overline{j}, j + 1] \to [0, 1]^2 \quad \text{and} \quad \rho_{2k+1} \circ \overline{T}_2 = T_2 \circ \rho_{2k+1} : A_g \to [0, 1]^n.
\]

Now we wish to apply Lemma 5.5 to the situation described in the following commutative diagram:

\[
\begin{array}{ccc}
\partial([0, 1]^2) & \xrightarrow{F} & [0, 1]^n \\
\rho_{2k+1} \downarrow & & \downarrow \rho_{2k+1} \\
[0, 1]^2 & \xrightarrow{f} & [0, 1]^n,
\end{array}
\]

where we define \(F := \overline{T}_2 \circ G \circ (\overline{T}_1)^{-1}\) and \(f := T_2 \circ g \circ (T_1)^{-1}\). To do so, we should confirm that the translations transfer to \(F\) the property of being the (restriction of the) standard cover of \(f\). To this end, consider \(F\) on the left-hand edge of the boundary \(\partial([0, 1]^2)\): for \(0 \leq t \leq 2k + 1\), we have
\[
F(\overline{0}, \overline{0} + t) = \overline{T}_2 \circ G(\overline{i}, \overline{j} + t) = \overline{T}_2(G(\overline{i}, \overline{j}) + t[G(i, j + 1) - g(i, j)])
\]
\[
= \overline{T}_2(g(\overline{i}, \overline{j})) + t[g(i, j + 1) - g(i, j)].
\]
Now, because $T_2$ is a translation, we may write
\[ g(i, j + 1) - g(i, j) = T_2 \circ g(i, j + 1) - T_2 \circ g(i, j) = f(0, 1) - f(0, 0), \]
and returning to the preceding expression we have
\[ F(0, 0 + t) = T_2(g(i, j)) + t[f(0, 1) - f(0, 0)] = f(0, 0) + t[f(0, 1) - f(0, 0)], \]
which is the correct expression for $F$ to be the standard cover of $f$ on the left-hand side of the square, as given in (21). The other sides are checked similarly; we omit those details.

So apply Lemma 5.5 to extend $F$ to $[0, 1]^2$, which we may then use to extend $G$ from the boundary of $[i, i + 1] \times [j, j + 1]$ to the map
\[ G := (T_2)^{-1} \circ F \circ T_1 : [i, i + 1] \times [j, j + 1] \to \overline{A_g}. \]

This extension fits into the following diagram, in which all internal parts other than the top trapezoid commute by construction, and in which we may reverse the directions of the translations and preserve commutativity. Since the outer square commutes by definition, it follows that the top trapezoid commutes too:

\[ \begin{array}{ccc}
[i, i + 1] \times [j, j + 1] & \xrightarrow{T_1} & [i, i + 1] \times [j, j + 1] \\
\xrightarrow{T_2} & \xrightarrow{[0, 1]^2} & \xrightarrow{[0, 1]^n}
\end{array} \]

The fact that the top trapezoid commutes means that, although $\overline{A_g}$ may contain points outside $S(Y, 2k + 1)$, nonetheless the image of the extended $G$ must be contained in $S(Y, 2k + 1)$, since the image of the original $g$ is contained in $Y$. That is, diagram (31) of the enunciation commutes, as claimed.

The last assertion follows from the construction of $F$ in Lemma 5.5. Referring to the notation used there, suppose we have $(s, t) \in [0, 1]^2$ in the central clique, so that $\{s, t\} \subseteq \{2k, 2k + 1\}$. Then the corresponding point $(s_b, t_b)$ is a corner point of the square, or $\{s_b, t_b\} \subseteq \{k, 3k + 1\}$, and we have $(s, t) = (C^k(s_b), C^k(t_b))$. Then $F$ is given by (see (22))
\[ F_i(s, t) = C^k(F_i(s_b, t_b)). \]
Since $F$ maps corners of the square to corners of the $n$-cube, we have $F_i(s_b, t_b) \in \{k, 3k + 1\}$ and so $C^k(F_i(s_b, t_b)) \in \{2k, kk + 1\}$. That is, each coordinate of $F(s, t)$ is either $2k$ or $2k + 1$ and $F(s, t)$ is a point in the central $2^n$-clique of $[\hat{\varnothing}, \hat{T}]^n$. Thus the image under $F$, as constructed in Lemma 5.5, of the central $4$-clique of $[\hat{\varnothing}, \hat{T}] \times [\hat{\varnothing}, \hat{T}]$ is contained in the central $2^n$-clique of $[\hat{\varnothing}, \hat{T}]^n$. As we observed in the discussion above Definition 5.1, translations $\overline{T}_c$ and $\overline{T}_2$ (and their inverses) preserve central cliques. The assertion about $G$ and central cliques follows. 

Now we consider the case in which the domain is a rectangle $I_M \times I_N$. The result here generalizes the results of the previous section in a very satisfactory way. Also, this case leads to a useful corollary about covers of homotopies (Corollary 6.2), which we use in [13].

**Theorem 5.7** Suppose we are given a map $H: I_M \times I_N \rightarrow Y$ for any digital image $Y \subseteq \mathbb{Z}^n$. For each integer $k \geq 1$, there is a canonical choice of map $\hat{H}: S(I_M \times I_N, 2k + 1) \rightarrow S(Y, 2k + 1)$ that makes the following diagram commute:

$$
\begin{array}{ccc}
S(I_M \times I_N, 2k + 1) & \xrightarrow{\hat{H}} & S(Y, 2k + 1) \\
\rho_{2k+1} & & \downarrow \rho_{2k+1} \\
I_M \times I_N & \xrightarrow{H} & Y.
\end{array}
$$

**Proof** Lemma 5.2 gives a map $\hat{H}: L_{M \times N} \rightarrow S(Y, 2k + 1)$ and we continue to extend this map over the whole of $S(I_M \times I_N, 2k + 1)$. We achieve this in two steps. Recall from (17) and (10) the definition of $\hat{H}$ in terms of the curves $\hat{\alpha}_j$ and $\hat{\beta}_j$.

**Extension Outside the Centers**

For $s < \hat{\varnothing} = k$ or $t < \hat{\varnothing} = k$, or $s > \hat{M} = (2k + 1)M + k$ or $t > \hat{N} = (2k + 1)N + k$ define

\[
\begin{align*}
\hat{H}(s, t) &= \hat{\beta}_0(t) & \text{for } 0 \leq s \leq k - 1 \text{ and } 0 \leq t \leq (2k + 1)N + 2k, \\
\hat{H}(s, t) &= \hat{\alpha}_0(s) & \text{for } 0 \leq s \leq (2k + 1)M + 2k \text{ and } 0 \leq t \leq k - 1, \\
\hat{H}(s, t) &= \hat{\beta}_M(t) & \text{for } \hat{M} + 1 \leq s \leq (2k + 1)M + 2k \text{ and } 0 \leq t \leq (2k + 1)N + 2k, \\
\hat{H}(s, t) &= \hat{\alpha}_N(s) & \text{for } 0 \leq s \leq (2k + 1)M + 2k \text{ and } \hat{N} + 1 \leq t \leq (2k + 1)N + 2k.
\end{align*}
\]

The situation now is illustrated in Fig. 10. Dots represent the points on which $\hat{H}$ has been defined at this point. Solid dots represent centers.

Notice that where definitions from this step overlap with each other, namely in each of the four corner regions, the definitions agree. For example, if $0 \leq s, t \leq k - 1$, we have $\hat{H}(s, t) = \hat{\beta}_0(t)$ and $\hat{H}(s, t) = \hat{\alpha}_0(s)$. Now, for $0 \leq t \leq k - 1$, Theorem 4.1 gives $\hat{\beta}_0(t) = \hat{\beta}_0(0) = \hat{H}(0, 0)$, and similarly we have $\hat{\alpha}_0(s) = \hat{H}(0, 0)$ for $0 \leq s \leq k - 1$. The other four corner regions behave similarly.
Now we check continuity, so far as we have defined $\hat{H}$. To this end, suppose we have adjacent points $(s, t)$ and $(s', t')$ in that part of $S(I_M \times I_N, 2k + 1)$ on which we have defined $\hat{H}$. If both points are in one of the horizontal bands $0 \leq t, t' \leq \bar{0} = k$ or $(2k + 1)N + k = \bar{N} \leq t, t' \leq (2k + 1)N + 2k$, then adjacency of $\hat{H}(s, t)$ and $\hat{H}(s', t')$ in $S(Y, 2k + 1)$ follows immediately from the continuity of the standard covers $\hat{\alpha}_j$. This is because, on these horizontal regions, we have defined $\hat{H}(s, t) = \hat{\alpha}_j(s)$, for a suitable $j = 0$ or $N$ depending on $t$. Hence, for $(s, t) \sim (s', t')$, we have $s \sim s'$ in $S(I_M, 2k + 1)$, whence $\hat{\alpha}_j(s) \sim \hat{\alpha}_j(s')$ and therefore $\hat{H}(s, t) \sim \hat{H}(s', t')$. For both points in one of the vertical bands $0 \leq s, s' \leq \bar{0} = k$ or $(2k + 1)M + k = \bar{M} \leq s, s' \leq (2k + 1)M + 2k$, adjacency of $\hat{H}(s, t)$ and $\hat{H}(s', t')$ in $S(Y, 2k + 1)$ follows immediately from the continuity of the standard covers $\hat{\beta}_i$, in a similar way. But if both adjacent points are not in either a horizontal or a vertical band of the type we have considered so far, then both must be in the lattice $L_{M \times N}$ considered in Lemma 5.2. Then we have $\hat{H}(s, t) \sim \hat{H}(s', t')$ from the continuity of $\hat{H}$ on $L_{M \times N}$ established in Lemma 5.2.

**Extension over Squares Whose Corners are Centers**

First note that we may extend $\hat{H}$ over the interior of any square in $S(I_M \times I_N, 2k + 1)$ whose corners are centers independently of any other such square. This is because any two points of $S(I_M \times I_N, 2k + 1)$ that are adjacent must be in one such square (including its edges) or, if not, then both must be in the region of $S(I_M \times I_N, 2k + 1)$ from the previous step, where we have already confirmed continuity. So it is sufficient to show that we may extend $\hat{H}$ over a typical such square $[i, i + 1] \times [j, j + 1]$ with corners

$$\{(i, j), (i + 1, j), (i, j + 1), (i + 1, j + 1)\}$$

for some $(i, j) \in I_M \times I_N$. Squares of this type are illustrated in Fig. 10—the eight empty squares in that figure. We defined $\hat{H}$ on the edges and corners of this square.
in Lemma 5.2. With reference to the definitions in (17) and (10), it is plain that \( \hat{H} \) is defined on \( \partial([i, i+1] \times [j, j+1]) \) as in Corollary 5.6. So applying Corollary 5.6 gives an extension over the whole square. Since we may extend \( \hat{H} \) over all such squares \([i, i+1] \times [j, j+1]\) one at a time, this completes the proof. \( \square \)

Before we extend Theorem 5.7 to the case in which the domain is an arbitrary 2D digital image, we make a simple observation.

**Remark 5.8** Whereas any digital image \( X \) may be contained in some sufficiently large rectangle, not every map of a digital image may be regarded as the restriction of some map of some containing rectangle. For example, take the identity map \( \text{id} : \partial(I_2 \times I_2) \to \partial(I_2 \times I_2) \). There is no (continuous) map \( I_2 \times I_2 \to \partial(I_2 \times I_2) \) that restricts to the identity on the boundary: the point \( (1, 1) \) is adjacent to every point of \( \partial(I_2 \times I_2) \), but there is no point of \( \partial(I_2 \times I_2) \) to which \( (1, 1) \) could be mapped and remain adjacent to every point on the boundary, which is what continuity would demand. We mention this to emphasize that it is not sufficient to consider maps of rectangles as in Theorem 5.7. This issue should be clear from a topological perspective: Our example corresponds to the well-known topological fact that a disc may not be retracted to its boundary; the issue is one of simple connectivity of the range. Making precise in the digital setting these kinds of topologically intuitive ideas is the business of our papers [13,14].

**Theorem 5.9** Suppose we are given a map \( f : X \to Y \) of digital images \( X \subseteq \mathbb{Z}^2 \) and \( Y \subseteq \mathbb{Z}^n \). For any integer \( k \geq 1 \), there is a map \( \hat{f} : S(X, 2k+1) \to S(Y, 2k+1) \) that makes the following diagram commute:

\[
\begin{array}{c}
S(X, 2k+1) \xrightarrow{\hat{f}} S(Y, 2k+1) \\
\rho_{2k+1} \downarrow \quad \downarrow \rho_{2k+1} \\
X \xrightarrow{f} Y
\end{array}
\]

**Proof** We use the notation established in previous results without comment. Begin by defining \( \hat{f} \) on centers, as

\[
\hat{f}(\vec{x}) = \overline{f(x)} \tag{33}
\]

for each \( x \in X \). Then, in each \( S(x, 2k+1) \subseteq S(X, 2k+1) \), extend \( \hat{f} \) horizontally and vertically from the center out to the edges of \( S(x, 2k+1) \) in one of two ways. For each vertical or horizontal neighbor \( x \leadsto x' \) in \( X \), interpolate the values of \( \hat{f} \) along the vertical or horizontal segment joining \( \vec{x} \) and \( \vec{x'} \) in \( S(X, 2k+1) \). Where \( x \) is missing its potential horizontal or vertical neighbor from \( X \), extend \( \hat{f} \) as a constant from the center out to that edge of \( S(x, 2k+1) \). In terms of a formula, suppose \( x \leadsto x' \) in \( \mathbb{Z}^2 \) (not necessarily in \( X \)), so that \( x' - x \) has one coordinate 0 and the other \( \pm 1 \). Then we define, for each \( t = 0, \ldots, k \),

\[
\hat{f}(\vec{x} + t(x'-x)) := \begin{cases} 
\overline{f(x)} + t[f(x') - f(x)] & x' \in X, \\
\overline{f(x)} & x' \notin X.
\end{cases} \tag{34}
\]
Fig. 11 \( \hat{f} \) defined on centers and extended horizontally and vertically within each \( S(x, 2k + 1) \) by (34). Illustrated with \( X = \{(0, 0), (2, 0), (1, 1), (2, 1), (2, 2)\} \) and \( 2k + 1 = 9 \)

Notice that, in case \( x' \in X \), we could equally well define

\[
\hat{f}(\bar{x} + t(x' - x)) = \hat{f}(x) + t[\hat{f}(x') - \hat{f}(x)] \quad \text{for } t = 0, \ldots, 2k + 1,
\]

to give values for \( \hat{f} \) on the segment in \( S(X, 2k + 1) \) that joins \( \bar{x} \) and \( \bar{x'} \) (including the endpoints), and this gives the same values as (34) on the relevant points of both \( S(x, 2k + 1) \) and \( S(x', 2k + 1) \). This amounts to using one of the formulas (30) of Corollary 5.6 along the side of a square between the centers \( \bar{x} \) and \( \bar{x'} \). This case of (34), in which \( x' \in X \), is how we proceeded in Theorem 5.7, when vertical and horizontal neighbors were always present except on the boundary of the rectangle. Notice also that, at this point, we do not interpolate in this way between diagonal neighbors of \( X \), such as \((x_1, x_2)\) and \((x_1 + 1, x_2 + 1)\).

See Fig. 11 for an illustration of the progress so far, for the case in which \( X \) consists of the five points \( X = \{(0, 0), (2, 0), (1, 1), (2, 1), (2, 2)\} \). As in the illustrations through the proof of Theorem 5.7, dots (open or closed) represent the points on which \( \hat{f} \) has been defined so far. Centers are represented as solid dots; we define \( \hat{f} \) on these points by (33). Open dots represent points on which we extend \( \hat{f} \) by (34).

Now we extend \( \hat{f} \) over the rest of \( S(X, 2k + 1) \). Since \( X \) is finite, we may pick some rectangle that contains it. Suppose we have \( X \subseteq [M_1, M_2] \times [N_1, N_2] \) for suitable
Fig. 12 $S(X, 2k + 1)$ covered by a rectangle $R = [M - 1, M + 1] \times [N - 1, N + 1]$. Centers of $R$ and squares of $R$ for which $I_{i, j} \neq \emptyset$ are illustrated with $2k + 1 = 9$.

$M_1, M_2$ and $N_1, N_2$. Then the squares

$[\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1] \quad \text{for} \quad M_1 - 1 \leq i \leq M_2 + 1, \quad N_1 - 1 \leq j \leq N_2 + 1$

cover the whole of $S(X, 2k + 1)$. Some of these squares may not include any points of $S(X, 2k + 1)$. But where $[\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1] \cap S(X, 2k + 1)$ is non-empty, we have already defined $\hat{f}$ on those parts of the boundary $\partial ([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1])$ that belong to $S(X, 2k + 1)$. We write $I_{i, j}$ for the intersection

$I_{i, j} := ([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]) \cap S(X, 2k + 1) \subseteq S(X, 2k + 1)$.

We will adapt Corollary 5.6 to extend $\hat{f}$ separately over each $I_{i, j}$. Conceptually, the squares $[\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]$ act as “cookie cutters,” dividing $S(X, 2k + 1)$, into various sub-regions over which we may extend $\hat{f}$ independently of each other. This latter observation holds for the same reason it held in the proof of Theorem 5.7: any two points adjacent in $S(X, 2k + 1)$ must both lie in a single region $I_{i, j}$ for some $i, j$. Thus, if we can extend $\hat{f}$ over each $I_{i, j}$ separately, and we already have $\hat{f}$ well-defined on their overlaps, then we may assemble the piecewise-defined map into a global, continuous $\hat{f}$ on the whole of $S(X, 2k + 1)$. In Fig. 12, we have illustrated the idea (same $X$ as in Fig. 11).
Here $S(X, 2k + 1)$ is included in the rectangle $[-1, 3] \times [-1, 3]$ and we have added (in grey) the centers $(\tilde{i}, \tilde{j})$ from $S([-1, 3] \times [-1, 3], 2k + 1)$. Also, we have indicated those squares (bounded by grey edges) $[\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]$ for which $I_{i,j}$ is non-empty. These $I_{i,j}$ consist of the union of any combination of the four quadrants of $[\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]$.

To extend $\tilde{f}$ over each $I_{i,j}$ we divide and conquer according as the number of quadrants of $[\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]$ included in $I_{i,j}$.

**$I_{i,j}$ Includes One Quadrant**

(This situation is illustrated by the square $[4, 13] \times [13, 22]$ in Fig. 12.) Exactly one corner of the square $[i, i + 1] \times [j, j + 1]$ lies in $X$; suppose it is $(c_1, c_2)$. Then in (33) we have already defined $\tilde{f}((c_1, c_2)) = f(c_1, c_2)$. Because $I_{i,j}$ includes only one quadrant, $(c_1, c_2)$ has no neighbors in $X$ that lie in $[i, i + 1] \times [j, j + 1]$, and (34) gives $\tilde{f}$ as constant at $f(c_1, c_2)$ on $\partial ([\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]) \cap S(X, 2k + 1)$. So in this case just extend $\tilde{f}$ to be constant at $f(c_1, c_2)$ on $I_{i,j}$. Obviously this give a continuous $\tilde{f}$ on $I_{i,j}$.

**$I_{i,j}$ Includes Two Quadrants**

(E.g. squares $[4, 13] \times [4, 13]$ or $[22, 31] \times [4, 13]$ in Fig. 12.) Exactly two corners of the square $[i, i + 1] \times [j, j + 1]$ lie in $X$. There are two sub-cases here, according as the corners are horizontal or vertical neighbors, or diagonal neighbors:

(a) One of the adjacent pairs

$$\{(i, j), (i, j + 1)\}, \quad \{(i + 1, j), (i + 1, j + 1)\},$$

$$\{(i, j), (i + 1, j)\}, \quad \{(i, j + 1), (i + 1, j + 1)\}$$

is in $X$ with the other two corners of $[i, i + 1] \times [j, j + 1]$ not in $X$;

(b) one of the adjacent pairs

$$\{(i, j), (i + 1, j + 1)\}, \quad \{(i + 1, j), (i, j + 1)\}$$

is in $X$ with the other two corners of $[i, i + 1] \times [j, j + 1]$ not in $X$.

For sub-case (5), assume the pair $\{(i, j), (i, j + 1)\} \subseteq X$ and that neither $(1 + 1, j)$ nor $(i + 1, j + 1)$ belongs to $X$. We will argue for this situation. It is easy to adapt our argument to handle the other three pairs; we leave their details for the reader. Begin by extending $f$ locally, and only for the purposes of taking this step, to a map

$$g: [i, i + 1] \times [j, j + 1] \rightarrow Y$$

by setting

$$g(i + 1, j) = g(i, j) = f(i, j) \quad \text{and} \quad g(i + 1, j + 1) = g(i, j + 1) = f(i, j + 1).$$
Now use \( g \) to define a map
\[
G : \partial ([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]) \to S(Y, 2k + 1)
\]
using the formulas (30) of Corollary 5.6. Because we have chosen \( g(i + 1, j) = g(i, j) \) and \( g(i + 1, j + 1) = g(i, j + 1) \), the formulas (30) give \( G \) constant on the top and bottom edges of \([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]\). As we remarked above when defining \( \hat{f} \) on that part of the boundary of \([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]\) included in \( I_{i,j} \), defining \( G \) by (30) on the left-hand edge agrees with how we defined \( \hat{f} \) here. So \( G \) agrees with \( \hat{f} \) where their definitions overlap. Now apply Corollary 5.6 to extend \( G \) over the whole of \([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]\), and then restrict \( G \) to \( I_{i,j} \) to obtain the desired extension of \( \hat{f} \). As the restriction of a continuous map, this extended \( \hat{f} \) is evidently continuous.

For sub-case (5), we again argue for just one situation and leave to the reader the (straightforward) details of adapting the argument to the other situation. Assume the pair \( \{(i, j), (i + 1, j + 1)\} \subseteq X \) and that neither \((1 + 1, j)\) nor \((i, j + 1)\) belongs to \( X \). Begin by extending \( f \) locally to a map
\[
g : [i, i + 1] \times [j, j + 1] \to Y
\]
by setting
\[
g(i + 1, j) = g(i, j + 1) = g(i, j) = f(i, j) \quad \text{and} \quad g(i + 1, j + 1) = f(i + 1, j + 1).
\]
As previously, use \( g \) to define a map \( G : \partial ([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]) \to S(Y, 2k + 1) \) using the formulas (30). Then apply Corollary 5.6 to extend \( G \) over the whole of \([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]\). Because we have chosen \( g(i + 1, j) = g(i, j + 1) = g(i, j) \), the formulas (30) give \( G \) constant on the left and bottom edges of \([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]\). Hence \( G \) and \( \hat{f} \) agree where their definitions overlap here. However, we may have \( g(i, j + 1) \neq g(i + 1, j + 1) \), in which case (30) would not give \( G \) constant on the top edge. Thus \( G \) and \( \hat{f} \) may not agree where their definitions overlap on the top edge. Similarly, \( G \) and \( \hat{f} \) may not agree on the right edge. So in this situation, restrict \( G \) to just the lower-left quadrant \([\bar{i}, \bar{i} + k] \times [\bar{j}, \bar{j} + k] \subseteq I_{i,j} \). Then this gives a continuous extension of \( \hat{f} \) over this lower-left quadrant. Repeat these steps for the upper-right quadrant. Namely, define \( g' : [i, i + 1] \times [j, j + 1] \to Y \) by setting
\[
g'(i, j) = f(i, j) \quad \text{and} \quad g'(i + 1, j) = g'(i, j + 1) = g'(i + 1, j + 1) = f(i + 1, j + 1).
\]
Extend over the boundary using (30) and then to \( G' \) over the whole of \([\bar{i}, \bar{i} + 1] \times [\bar{j}, \bar{j} + 1]\) using Corollary 5.6. Whereas \( G' \) may not agree with \( \hat{f} \) on the boundary of the lower-left quadrant, it does agree with \( \hat{f} \) on the boundary of the upper-right. Hence, restricting \( G' \) to the upper-right quadrant \([\bar{i} + k + 1, \bar{i} + 1] \times [\bar{j} + k + 1, \bar{j} + 1] \subseteq I_{i,j} \) gives a continuous extension of \( \hat{f} \) over this upper-right quadrant. There remains the technical point of whether these two continuous extensions on lower-left and upper-right quadrants assemble into a continuous map on \( I_{i,j} \). The issue here is whether we
have \( G(\tilde{i} + k, \tilde{j} + k) \sim_Y G'(\tilde{i} + k + 1, \tilde{j} + k + 1) \), since the points \((\tilde{i} + k, \tilde{j} + k)\) (in the lower-left quadrant) and \((\tilde{i} + k + 1, \tilde{j} + k + 1)\) (in the upper-right) are the only pair of adjacent points of \(I_{i,j}\) that do not lie in a single quadrant. Now the final assertion of Corollary 5.6 gives that \(G(\tilde{i} + k, \tilde{j} + k)\) is some point in the central \(2^n\)-clique of the cube \(\tilde{A}_g\). It also gives that \(G'(\tilde{i} + k + 1, \tilde{j} + k + 1)\) is some point in the central \(2^n\)-clique of the cube \(\tilde{A}_{g'}\). The key point here, though, is that because we gave \(g\) and \(g'\) values at each corners of \([\tilde{i}, \tilde{j} + 1] \times [\tilde{j}, \tilde{j} + 1]\) chosen from the same set, namely, \(\{f(i, j), f(i + 1, j + 1)\}\), it follows that we have identical cubes \(A_g = A_{g'}\) and thus \(\tilde{A}_g = \tilde{A}_{g'}\). Then \(G(\tilde{i} + k, \tilde{j} + k)\) and \(G'(\tilde{i} + k + 1, \tilde{j} + k + 1)\) are points in the central \(2^n\)-clique of the same cube, and thus are adjacent. It follows that we have our continuous extension of \(\hat{f}\) over \(I_{i,j}\).

**\(I_{i,j}\) Includes Three Quadrants**

(E.g. square \([13, 22] \times [4, 13]\) in Fig. 12.) Exactly three corners of the square \([i, i + 1] \times [j, j + 1]\) lie in \(X\). Once again we argue for just one situation and leave to the reader the (straightforward) details of adapting the argument to the other three situations. Assume we have \(\{(i, j), (i + 1, j), (i, j + 1)\} \subseteq X\) and that \((i + 1, j + 1)\) does not belong to \(X\). That is, the upper-right quadrant is missing from \(I_{i,j}\). First extend \(f\) locally to a map \(g: [i, i + 1] \times [j, j + 1] \to Y\) by setting

\[
g(i, j) = f(i, j), \quad g(i, j + 1) = f(i, j + 1), \quad \text{and} \quad g(i + 1, j + 1) = g(i + 1, j) = f(i + 1, j).
\]

With \(g\) define \(G: \partial([\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]) \to S(Y, 2k + 1)\) using the formulas (30). Then apply Corollary 5.6 to extend \(G\) over the whole of \([\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]\). Because we have chosen \(g(i + 1, j + 1) = g(i + 1, j)\), the formulas (30) give \(G\) constant on the right edge of \([\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]\), agreeing with \(\hat{f}\) where their definitions overlap on this edge. Also, since \(g\) and \(f\) agree on the three corners of \([i, i + 1] \times [j, j + 1]\) that lie in \(X\), then \(G\) and \(\hat{f}\) agree on the left and bottom edges of \([\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]\). However, we may have \(g(i, j + 1) \neq g(i + 1, j + 1)\), in which case (30) would not give \(G\) constant on the top edge. Thus \(G\) and \(f\) may not agree where their definitions overlap on the top edge. So here, restrict \(G\) to just the lower two quadrants \([\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + k] \subseteq I_{i,j}\). Then this gives a continuous extension of \(\hat{f}\) over these lower two quadrants. Repeat these steps for the left two quadrants. Namely, define \(g': [i, i + 1] \times [j, j + 1] \to Y\) by setting

\[
g'(i, j) = f(i, j), \quad g'(i + 1, j) = f(i + 1, j), \quad \text{and} \quad g'(i + 1, j + 1) = g'(i, j + 1) = f(i, j + 1).
\]

Extend over the boundary using (30) and then to \(G'\) over the whole of \([\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]\) using Corollary 5.6. Whereas \(G'\) may not agree with \(\hat{f}\) on the right edge of the lower-right quadrant, it does agree with \(\hat{f}\) on the top edge of the upper-left, and also on the left and bottom edges of \([\tilde{i}, \tilde{i} + 1] \times [\tilde{j}, \tilde{j} + 1]\). Hence, restricting \(G'\) to the
left two quadrants $[i, i + k] \times [j, j + 1] \subseteq I_{i,j}$ gives a continuous extension of $\hat{f}$ over these left two quadrants. Furthermore, because $G$ and $G'$ agree on the left and bottom edges of the lower-right quadrant, their extensions agree on this lower-right quadrant. For this last assertion, note that we are relying on the cubes $A_g$ and $A_{g'}$, and thus $\overline{A_g}$ and $\overline{A_{g'}}$, being the same as each other. As in the previous case, this follows because we assigned $g$ and $g'$ values at each corner of $[i, i + 1] \times [j, j + 1]$ chosen from the same set, namely, $(f(i, j), f(i + 1, j), f(i, j + 1))$. Once we have $A_g = A_{g'}$, it is obvious that $G$ and $G'$ agree on the lower-right quadrant because they agree on the left and bottom edges of the lower-right quadrant, and their extensions are determined as in Corollary 5.6 by their values on the boundary once the cube $\overline{A_g} = \overline{A_{g'}}$ is determined. Thus far, we have a well-defined extension of $\hat{f}$ over $I_{i,j}$, given by $G$ on the lower two quadrants and $G'$ on the left two quadrants. This is a continuous map so long as we have $G'(i + k + 1, j + k) \sim Y G''(i + k, j + k + 1)$, since the points $(i + k + 1, j + k)$ (in the lower-right quadrant) and $(i + k, j + k + 1)$ (in the upper-left) are the only pair of adjacent points of $I_{i,j}$ that do not both lie in either the lower two quadrants or in the left two quadrants. As in the previous case, $G(i + k + 1, j + k)$ and $G'(i + k, j + k + 1)$ are points in the central $2^n$-clique of the same cube $\overline{A_g} = \overline{A_{g'}}$, and thus are adjacent. It follows that we have our continuous extension of $\hat{f}$ over $I_{i,j}$. The last case, in which $I_{i,j}$ includes all four quadrants of $[i, i + 1] \times [j, j + 1]$, is simply Corollary 5.6. We have exhausted all the possible cases and obtained an extension of $\hat{f}$ over $I_{i,j}$ for which the diagram

$$
\begin{array}{ccc}
I_{i,j} & \xrightarrow{\hat{f}} & S(Y, 2k + 1) \\
\downarrow{\rho_{2k+1}} & & \downarrow{\rho_{2k+1}} \\
([i, i + 1] \times [j, j + 1]) \cap X & \xrightarrow{f} & Y
\end{array}
$$

commutes. As we explained earlier, these maps on each $I_{i,j}$ may be assembled into a continuous $\hat{f}$ on the whole of $X$. □

6 Covering of Homotopies for Paths and Loops; Future Work

We give a consequence of Theorem 5.7 that, together with Theorem 4.1, provides results similar to path lifting and homotopy lifting results that play a prominent role in the development of the fundamental group in the ordinary topological setting. In fact we rely on this result in [13], where we develop a digital fundamental group.

We use a “cylinder object” definition of homotopy, which is the one commonly used in the digital topology literature. In [14] we give a fuller discussion of homotopy, including a “path object” definition as well.

**Definition 6.1** Let $f, g : X \rightarrow Y$ be (continuous) maps of digital images. We say that $f$ and $g$ are homotopic, and write $f \approx g$, if, for some $N \geq 1$, there is a continuous
map

\[ H : X \times I_N \to Y, \]

with \( H(x, 0) = f(x) \) and \( H(x, N) = g(x) \). Then \( H \) is a homotopy from \( f \) to \( g \).

In Theorem 5.7 the map \( H : I_M \times I_N \to Y \) has domain a rectangle and may be viewed as a homotopy of paths in \( Y \). Since the subdivided rectangle is again a (larger) rectangle, the map \( \hat{H} \) that covers \( H \) in Theorem 5.7 may also be viewed as a homotopy of paths but now in \( S(Y, 2k + 1) \). Namely, \( \hat{H} \) is a homotopy that covers the homotopy \( H \) and Theorem 5.7 may be seen as a covering homotopy theorem to accompany the covering path theorem of Theorem 4.1.

Suppose we have paths (of the same length) in \( Y \) with the same initial and terminal points. That is, we have maps \( \alpha, \beta : I_M \to Y \) with \( \alpha(0) = \beta(0) = y_0 \) and \( \alpha(M) = \beta(M) = y_M \) for some \( y_0, y_M \in Y \). If \( \alpha \approx \beta \), then the homotopy may be relative the endpoints, which is to say that we have \( H(0, t) = y_0 \) and \( H(M, t) = y_M \) for all \( t \in I_N \). If \( \alpha \) and \( \beta \) are loops in \( Y \), so that \( y_M = y_0 \), and if \( \alpha \approx \beta \) via a homotopy relative the endpoints, then we say that \( \alpha \) and \( \beta \) are homotopic via a based homotopy of based loops. The nomenclature comes from the setting of the fundamental group, as in [13], in which \( Y \) is a based digital image, and maps, loops, and homotopies are based.

The construction of \( \hat{H} \) in the proof of Theorem 5.7 leads to the following “covering homotopy” property of subdivisions.

**Corollary 6.2** (to Theorem 5.7) Suppose \( \alpha, \beta : I_M \to Y \) are paths in \( Y \), with standard covers \( \hat{\alpha}, \hat{\beta} : S(I_M, 2k + 1) = I(2k + 1)M + 2k \to S(Y, 2k + 1) \) as in Theorem 4.1.

(A) If \( \alpha \approx \beta \), then \( \hat{\alpha} \approx \hat{\beta} : S(I_M, 2k + 1) \to S(Y, 2k + 1) \).

(B) Suppose we have \( \alpha(0) = \beta(0) = y_0 \) and \( \alpha(M) = \beta(M) = y_M \) for some \( y_0, y_M \in Y \). If \( \alpha \approx \beta \) relative the endpoints, then \( \hat{\alpha} \approx \hat{\beta} \) relative the endpoints.

(C) Suppose we have \( \alpha(0) = \alpha(M) = y_0 = \beta(0) = \beta(M) \), so that \( \alpha \) and \( \beta \) are loops based at some \( y_0 \in Y \). If \( \alpha \approx \beta \) via a based homotopy of based loops, then \( \hat{\alpha} \approx \hat{\beta} \) via a based homotopy of based loops.

**Proof** Part (A) is more-or-less a re-statement of the behavior of \( \hat{H} \) around the edges of the rectangle, from Theorem 5.7. It follows from the construction of \( \hat{H} \). In part (B), notice that we have \( \hat{\alpha}(0) = \hat{\beta}(0) = \overline{y_0} \) and \( \hat{\alpha}((2k+1)M+2k) = \hat{\beta}((2k+1)M+2k) = \overline{y_M} \) from the construction of Theorem 4.1. Then part (B) follows from the construction of \( \hat{H} \) in Theorem 5.7, together with the fact that the standard cover of a constant path is a constant path (part (a) of Lemma 4.4). In part (C), we have that \( \hat{\alpha} \) and \( \hat{\beta} \) are loops in \( S(Y, 2k + 1) \) (of length \( (2k + 1)M + 2k \)) based at \( \overline{y_0} \) from the construction of Theorem 4.1. Then part (C) is a special case of part (B). \( \square \)

The ability to cover based homotopies of based loops in this way leads in [13] to the result that our fundamental group constructed there is preserved by subdivision. That result is one of the major advances of [13] over existing treatments of the fundamental group in the digital topology literature. Other applications of the results of this paper appear in [14].
The results of this paper involve odd subdivisions and maps with 1D and 2D domains. We believe that comparable results may be given for even subdivisions; their statements and proofs would involve an adaptation to the fact that we have no middle point in an interval of odd length. Furthermore, we believe our results here may be extended to treat maps with domains of any dimension. Doing so would allow the approach of [13] to be extended to a comparable development of higher homotopy groups, for instance. To extend our results in these ways, however, involves resolving many technical details as well as expositional challenges. We hope to make progress in this direction in future work.

References

1. Andres, E.: Discrete circles, rings and spheres. Comput. Graph. 18(5), 695–706 (1994)
2. Barrera, T., Hast, A., Bengtsson, E.: A chronological and mathematical overview of digital circle generation algorithms—introducing efficient 4- and 8-connected circles. Int. J. Comput. Math. 93(8), 1241–1253 (2016)
3. Borat, A., Vergili, T.: Digital Lusternik–Schnirelmann category. Turkish J. Math. 42(4), 1845–1852 (2018)
4. Boxer, L.: A classical construction for the digital fundamental group. J. Math. Imaging Vis. 10(1), 51–62 (1999)
5. Boxer, L.: Remarks on digitally continuous multivalued functions. J. Adv. Math. 9(1), 1755–1762 (2014)
6. Boxer, L., Staecker, P.C.: Connectivity preserving multivalued functions in digital topology. J. Math. Imaging Vis. 55(3), 370–377 (2016)
7. Escribano, C., Giraldo, A., Sastre, M.A.: Digitally continuous multivalued functions. In: Discrete Geometry for Computer Imagery (Lyon 2008). Lecture Notes in Comput. Sci., vol. 4992, pp. 81–92. Springer, Berlin (2008)
8. Escribano, C., Giraldo, A., Sastre, M.A.: Digitally continuous multivalued functions, morphological operations and thinning algorithms. J. Math. Imaging Vis. 42(1), 76–91 (2012)
9. Evako, A.V.: Topological properties of closed digital spaces: one method of constructing digital models of closed continuous surfaces by using covers. Comput. Vis. Image Underst. 102(2), 134–144 (2006)
10. Evako, A.V.: Classification of digital n-manifolds. Discrete Appl. Math. 181, 289–296 (2015)
11. Farkas, J., Bajká, Sz., Nagy, B.: Approximating the Euclidean circle in the square grid using neighbourhood sequences. Pure Math. Appl. 17(3–4), 309–322 (2006)
12. Kim, C.E.: Digital disks. IEEE Trans. Pattern Anal. Mach. Intell. 6(3), 372–374 (1984)
13. Lupton, G., Oprea, J., Scoville, N.A.: A fundamental group for digital images. J. Appl. Comput. Topol. 5(2), 249–311 (2021)
14. Lupton, G., Oprea, J., Scoville, N.A.: Homotopy theory in digital topology. Discrete Comput. Geom. 67(1), 112–165 (2022)
15. Rosenfeld, A.: ‘Continuous’ functions on digital pictures. Pattern Recognit. Lett. 4(3), 177–184 (1986)
16. Rosenfeld, A., Pfaltz, J.L.: Distance functions on digital pictures. Pattern Recognit. 1(1), 33–61 (1968)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.