

Cycle construction and geodesic cycles with application to the hypercube

Paul C. Kainen

*Department of Mathematics and Statistics
Georgetown University, Washington, DC 20057 USA*

Received 2 March 2013, accepted 23 April 2014, published online 9 June 2014

Abstract

Construction of cycles in a graph is investigated, where cycles from particular subsets (such as bases) are added together so that each partial sum is also a cycle or each new cycle intersects the sum of the preceding terms in a nontrivial path. Starting with the geodesic cycles, a hierarchical construction is given. For the hypercube graph, geodesic cycles are characterized, and it is shown how hypercube geodesic cycles can be constructed in two steps from a special basis. Applications are given to inferring commutativity of a diagram in a groupoid from commutativity of some of its cycles.

Keywords: Robust cycle basis, well-arranged sequence, geodesic cycle, Cayley graph, hypercube, forcing commutativity, groupoid diagram.

Math. Subj. Class.: 05C38, 18B40

1 Introduction

Let G be a graph. A cycle-subgraph of G is a connected, 2-regular subgraph. Let $Cyc(G)$ denote the set of all cycle-subgraphs of G . Call $\mathcal{S} \subseteq Cyc(G)$ “weakly robust” if for every $z \in Cyc(G)$, there is an integer $k \geq 1$ and $z_1, \dots, z_k \in \mathcal{S}$ (not necessarily distinct) such that

$$z = z_1 + z_2 + \cdots + z_k \tag{1.1}$$

and for $2 \leq j \leq k - 1$

$$z_1 + z_2 + \cdots + z_j \in Cyc(G); \tag{1.2}$$

that is, \mathcal{S} is weakly robust if it spans the cycle space of G in such a way that every cycle-subgraph has all partial sums remaining cycle-subgraphs (not just Eulerian graphs). Call \mathcal{S} “robust” if it is weakly robust and if, in addition, for $2 \leq j \leq k$,

$$z_j \cap (z_1 + \cdots + z_{j-1}) \text{ is a nontrivial path.} \tag{1.3}$$

E-mail address: kainen@georgetown.edu (Paul C. Kainen)

Of course, $\mathcal{S} = Cyc(G)$ has these properties vacuously, taking $k = 1$. Also, (1.3) \implies (1.2).

The problem of finding a cycle basis which is weakly robust seems to have first appeared in print in a paper of Dixon and Goodman [2]. Given a spanning tree in a connected graph, one can construct a basis of cycles using the non-tree edges. In [2], it was conjectured that such bases are weakly robust but this was disproved by Sysło [22]. Later, it was shown by Doğrusöz and Krishnamoorthy [3] that the set of region boundaries of a plane graph, excluding the unbounded region, constitute a weakly robust basis.

The problem was rediscovered in [12], including the new question of finding a robust basis, where we showed that for complete graphs, the set of all 3-cycles containing a fixed vertex determines a robust basis. A basis which is not weakly robust (due to A. Vogt) was exhibited. We also found a basis for the bipartite complete graph (the set of all 4-cycles through a fixed edge), which was shown to be weakly robust by Ostermeier et al. [19]. Our motivation was category theory and graph theory, while the work of Stadler and co-authors [9, 11, 15, 16, 19, 8] is also motivated by applications, especially in biochemistry. These latter papers, as well as [12], use notation which is different from that used here.

The properties of cycle bases and related sets discussed below were developed to permit the propagation of commutativity from a basis of cycles to all of the diagram; see [13], [14]. In this paper, it is shown that the commutativity of a d -dimensional hypercube diagram in a groupoid can be *forced* by the commutativity of a carefully chosen subset of the square faces, containing approximately a fraction of $4/d$ of the square faces. In contrast, [13] showed that commutativity cannot be *blocked* by any family with fewer than $d - 2$ squares in a d -cube; that is, if all but $d - 2$ squares in a d -cube diagram in a groupoid are known to commute, then the entire diagram is commutative.

The paper is organized as follows: Section 2 gives some basic graph-theoretic, algebraic, and topological background. The next section reviews the basic types of well-arranged sequences of cycles, providing the machinery used in the remaining sections. Section 4 covers robustness and weak robustness, including the idea of iterating these constructions, and relates them to Cayley color graphs. In section 5 we give a brief proof that the basis of region boundaries in a plane graph is robust. Section 6 shows how geodesic cycles can build all cycles through an iterative process. Sections 7 and 8 apply only to the hypercube. Geodesic cycles in the d -cube are given a simple characterization and then they are constructed from a recursively defined basis. Section 9 derives the application to commutative diagrams in groupoids and shows that a square which is the sum of two commutative squares can be non-commutative if the two squares added do not intersect in a nontrivial path. Finally, in Section 10, other types of cycle bases are considered.

2 Basic concepts

In this paper, graphs are finite, undirected, and simple (i.e., no loops or parallel edges). For basic graph theory, see, e.g., [7], [1]. A connected graph is *Eulerian* if all of its vertex-degrees are even. For G a graph, $V(G)$ and $E(G)$ denote the sets of vertices and edges, respectively, and we also write $G = (V, E)$. Every graph has a representation as a family of points and closed straight-line segments (corresponding to the vertices and edges, resp.) embedded in 3-dimensional Euclidean space; each segment in the representation joins the two points corresponding to the edge which the segment represents. There is a unique topology which induces the usual topology of a closed interval on the straight-line

segments, and this is called the *underlying topology* of the graph. A *cycle* is a graph which is 2-regular and connected. The underlying topology of a cycle is the unique closed connected 1-manifold, i.e., the circle. Let G be a graph and let $Cyc(G)$ denote the set of all subgraphs of G which are cycles.

A *path* is a graph which consists of a single vertex, or a connected graph which contains two vertices of degree 1 with every other vertex of degree 2. A nontrivial path is a path which contains more than one vertex. We write P_r for the path with r vertices.

Let \mathbf{Z}_2 denote the field with 2 elements $\{0, 1\}$; \mathbf{Z}_2^A denotes the \mathbf{Z}_2 -vector space of all functions from a set A to \mathbf{Z}_2 ; $\dim \mathbf{Z}_2^A = |A|$, the cardinality of A . Addition of functions is just coordinate-wise addition (modulo 2), so the zero element is the constant 0 function. To each function in \mathbf{Z}_2^A , one can associate the subset of A determined by those elements mapping to 1 and conversely for every subset of A , there is a unique element in \mathbf{Z}_2^A to which it is associated (the “characteristic function” of the set). The symmetric difference of two subsets of A corresponds to the vector space sum of the corresponding elements in \mathbf{Z}_2^A [1, p. 23]. If U is any subset of a vector space V , then $span(U)$ is defined to be the intersection of all linear subspaces of V which contain U . Over the field \mathbf{Z}_2 , the linear span of U is just the set of all sums of subsets of U ; if U is empty, its span is 0.

If $G = (V, E)$ is a graph, there is a linear map

$$\partial_G : \mathbf{Z}_2^{E(G)} \rightarrow \mathbf{Z}_2^{V(G)}$$

given by $\partial_G(e) = v + w$ when $e = vw \in E(G)$. The members of $\mathbf{Z}_2^{E(G)}$ are called the 1-chains while members of $\mathbf{Z}_2^{V(G)}$ are called 0-chains. The kernel (or “null space”) $Z(G)$ of the \mathbf{Z}_2 -linear map ∂_G is the subset of the 1-chains which map to the zero 0-chain. Thus, connected components of subgraphs of G induced by elements of $Z(G)$ correspond to Eulerian subgraphs of G . By a theorem of Euler, H in $Z(G)$ is an edge-disjoint union, hence sum, of cycles [1, p. 24], so $Z(G)$ is the linear span of $Cyc(G)$. So there is a basis for $Z(G)$ consisting entirely of cycles, which is called a *cycle basis*. Addition in $Z(G)$ is denoted $+$. One refers to the vector space $Z(G)$ as the “cycle space” of the graph where the word “cycle” is used in the algebraic, rather than geometric, sense.

The *cyclomatic number* $b(G)$ of a graph G is the dimension of $Z(G)$. It is well-known (and straightforward to show) that

$$b(G) = q - p + k,$$

where q, p, k denote the number of edges, vertices, and connected components of G .

If G is connected, then for every spanning tree T of G , there is a basis $\mathcal{B}(G, T)$ which is the set of cycles z_e formed by adding a single element $e = vw \in E(G) \setminus E(T)$. Indeed, as each pair v, w of distinct vertices in G are joined by a unique path in T , $T + e$ has a unique cycle. Each of these cycles has an edge contained by no other cycle, so the set of cycles must be independent. But the cardinality of this independent set is maximal, for T contains $p - 1$ edges, so $|\mathcal{B}(G, T)| = q - p + 1$, and so it is a basis, the *spanning tree basis*,

3 Well-arranged sequences of cycles

If $z, z' \in Cyc(G)$, we write $z||z'$ and call z, z' *compatible* if $z \cap z' = P$, where P is a nontrivial path. This relation on the set of cycles is symmetric and irreflexive.

For two compatible cycles z, z' , the underlying topology of $z + z'$ is the connected sum of the underlying topologies of the components, so $z + z'$ must be a cycle as well. We

give a direct argument avoiding the reference to connected sum after introducing a useful notation.

If z is any cycle and P is a nontrivial path contained in z , then there is a unique non-trivial path P' also contained in z such that $z = P \cup P'$ and $P \cap P' = 2K_1$ (i.e., P and P' intersect in two isolated vertices). We write $z - P$ for P' .

Lemma 3.1. *The sum of two compatible cycles is a cycle.*

Proof. Let z, z' be two cycles with $z \cap z' = P$ a nontrivial path. Then

$$z + z' = (z - P) \cup (z' - P).$$

The paths $z - P$ and $z' - P$ intersect in the two endpoints of P , so their union is a cycle. \square

Let $z \in Cyc(G)$; suppose $\mathcal{S} \subseteq Cyc(G)$ and $span(\mathcal{S}) = Z(G)$; e.g., \mathcal{S} could be a cycle basis. A z -sequence in \mathcal{S} is a sequence $\sigma = (s_1, s_2, \dots, s_k) \in \mathcal{S}^k$ such that

$$z = s_1 + s_2 + \dots + s_k = \sum_{i=1}^k s_i.$$

We do not insist that the entries in σ need all be distinct. Of course, for every cycle z there are z -sequences in \mathcal{S} with all members distinct since \mathcal{S} spans every cycle, but in order to achieve further constraints on the sequence, it may be necessary to allow repetitions of members of the sequence. When there are no repetitions, it will be indicated.

For a sequence $\sigma = (s_1, s_2, \dots, s_k) \in \mathcal{S}^k$, there are several notions of “well-arranged”: σ is *well-arranged with respect to intersection (wai)* if for each $j, 2 \leq j \leq k$, s_j intersects the previous partial sum $\sum_{i=1}^{j-1} s_i$ in a nontrivial path, that is, $s_j \parallel \left(\sum_{i=1}^{j-1} s_i\right)$ for $2 \leq j \leq k$;

σ is *well-arranged with respect to topology (wat)* if each partial sum is a cycle; that is, if the partial sums are topologically constant. These are the only variants of well-arrangedness which we shall need below but for completeness we mention two more.

σ is *well-arranged with respect to connectedness (wac)* if each partial sum is connected;

σ is *well-arranged with respect to degree (wad)* if each partial sum is regular of degree 2.

By Lemma 3.1, **wai** \implies **wat**. However, the converse is false. Noncompatible cycles may sum to a cycle (see, e.g., the example at the end of section 6), so there can be z -sequences which are well-arranged with respect to topology but not with respect to intersection.

Note that **wat** \implies **wac** and **wat** \implies **wad** trivially. Conversely, if **wad** holds for the z -sequence σ , then σ satisfies a weaker intersection condition that for each $j, 2 \leq j \leq k$, s_j intersects the partial sum $\sum_{i=1}^{j-1} s_i$ in a subgraph with no isolated vertices. Further, this weak intersection condition, in turn, implies **wad**.

4 Robustness and weak robustness

Let G be a graph and let $\mathcal{S} \subseteq Cyc(G)$. We define the *robust closure* $\rho(\mathcal{S})$ (or *weak robust closure* $\rho_w(\mathcal{S})$) of \mathcal{S} with respect to G to be the set of all cycles in G which are the sum of a **wai** (resp., **wat**) sequence of elements in \mathcal{S} . A set \mathcal{S} of cycles is called *robust* (*weakly robust*) if $\rho(\mathcal{S}) = Cyc(G)$ (resp. $\rho_w(\mathcal{S}) = Cyc(G)$). Of course, a weakly robust (and

so also a robust) set of cycles is necessarily spanning: $\text{span}(\mathcal{S}) = Z(G)$. When \mathcal{S} is a basis for $Z(G)$, then every cycle can be written as the sum of a unique subset of the basis, with no element repeated. This non-repetition property of the sequence is not required for the definition of robust and weak robust closure. However, no examples are known where every **wai** (or **wat**) sequence which sums to some cycle z has repeated terms.

Hierarchical iteration is achieved by taking the robust span of the set of cycles produced by the previous stage. Let $\rho^0(\mathcal{S}) = \mathcal{S}$ and, for integer $a \geq 1$, put $\rho^a(\mathcal{S}) := \rho(\rho^{a-1}(\mathcal{S}))$. Then $\rho^a(\mathcal{S}) \subseteq \rho^b(\mathcal{S})$ when $a \leq b$, for a, b nonnegative integers. Also, if $\mathcal{S} \subseteq \mathcal{T}$, then $\rho(\mathcal{S}) \subseteq \rho(\mathcal{T})$.

We note the following for use in Section 9. Let $\mathcal{T} \subseteq \text{Cyc}(G)$ for some graph G . Call \mathcal{T} *cooperative* if $(z, z' \in \mathcal{T} \text{ and } z \parallel z') \implies z + z' \in \mathcal{T}$. According to Lemma 3.1, the set of all cycles is cooperative. From the definition of iterated robust closure and our remarks there, if G is any graph, one has the following.

Lemma 4.1. *If $\mathcal{S} \subseteq \mathcal{T} \subseteq \text{Cyc}(G)$, \mathcal{T} is cooperative, and k is a nonnegative integer, then*

$$\rho^k(\mathcal{S}) \subseteq \mathcal{T}.$$

Hierarchically iterated robust closure and the lemma above are applied in Cor. 9.3 to obtain results on the propagation of commutativity in general groupoid diagrams, especially when the diagram has the scheme of a hypercube. (Definitions are given there.)

These results can be expressed in terms of Cayley color graphs [24]; see [16], [19], where the concept is used but not the name. Recall that the *Cayley color graph* of a group \mathcal{A} with respect to a subset $\mathcal{S} \subseteq \mathcal{A}$ is the digraph $\Gamma(\mathcal{A}, \mathcal{S})$ with \mathcal{A} as its vertex-set, where for vertices v, w , there is an arc $a = (v, w)$ if and only if there exists an element $u \in \mathcal{S}$ such that $w = v + u$. (We write “+” as we are only considering the Abelian groups underlying vector spaces.) The arc $a = (v, w)$ is *colored* by assigning to it the unique u satisfying the equation $w = v + u$, and so the Cayley color graph is the ordered pair consisting of the digraph and the corresponding arc colors.

Let \mathcal{A} be the abelian group determined by $Z(G)$. Each element has order 2, so the Cayley color graph is symmetric and can be replaced by an undirected graph. To avoid loops, one assumes that the identity element of the group is not in the set \mathcal{S} . Plainly, \mathcal{S} is spanning if and only if $\Gamma(\mathcal{A}, \mathcal{S})$ is connected; see, e.g., [19, Lemma 3.3].

Let G be a graph, let $\mathcal{S} \subseteq \text{Cyc}(G)$ with $\text{span}(\mathcal{S}) = Z(G)$. Let $\Gamma_{\text{Cyc}}(G, \mathcal{S})$ denote the subgraph of $\Gamma(\mathcal{A}, \mathcal{S})$ induced by $\text{Cyc}(G) \subseteq \mathcal{A}$. Given two cycles z, z' in G there is a z' - z -path in $\Gamma_{\text{Cyc}}(G, \mathcal{S})$ joining them if and only if there is a **wat** z -sequence σ in \mathcal{S} and such that $z' = s_1$. It is convenient to take z' to be the null-set of edges which is 0 in the vector space. The set of all cycles z reachable from 0 by such paths is exactly the weak robust span of \mathcal{S} . The graph $\Gamma_{\text{Cyc}}(\mathcal{A}, \mathcal{S})$ is connected if and only if there is a $\Gamma_{\text{Cyc}}(\mathcal{A}, \mathcal{S})$ -path joining 0 to each $z \in \text{Cyc}(G)$, that is, if and only if \mathcal{S} is weakly robust.

The robust span of \mathcal{S} is the set of all vertices in the connected component of 0 in the edge-induced subgraph $\Gamma'_{\text{Cyc}}(G, \mathcal{S})$ of $\Gamma_{\text{Cyc}}(G, \mathcal{S})$ determined by those edges zz' where $z \parallel z'$. In this case, the partial sums are successively modified by replacing a single subpath P with a complementary path $P' = z - P$, where $P \subset z$, for z some member of \mathcal{S} .

Instead of robustness or weak robustness, given any family \mathcal{S} of cycles which spans $Z(G)$, one might investigate the least number of connected components, the least maximum degree, or the least number of distinct topologies which occur along the $\Gamma_{\text{Cyc}}(\mathcal{A}, \mathcal{S})$ (or

$\Gamma'_{Cyc}(\mathcal{A}, \mathcal{S})$ paths from 0 to z , maximized or totaled over all cycles z in G . One may also focus only on some cycles which need to be built up. See [15] for an application to random networks and also Proposition 10.1 below.

5 Planarity and robustness

For plane graphs, we find a robust basis; cf. [3]. In the following, we do not distinguish between subgraphs of a plane graph and the corresponding subsets of the plane. Let G be a 2-connected plane graph. To prevent pathological embeddings, assume that each edge of G is embedded as a piecewise-linear curve (with a finite number of segments). Then $\mathbf{R}^2 \setminus G$ is a disjoint union of open disks and, using the 2-connectedness of the graph, the closure of each open disk is a closed disk and the boundary of each closed disk is a cycle.

Each plane graph contains exactly one unbounded region, containing points which are arbitrarily far from any point of the graph; all the other regions are called *bounded*. See, e.g., [1, Chap. 4]. Let $\mathcal{R}(G)$ denote the set of all bounded regions. Then

$$\mathcal{B}(G) := \{\partial r : r \in \mathcal{R}(G)\}$$

is a basis of the cycle space of G . It is spanning since if $z \in Cyc(G)$, then $z = \sum \partial r$, summing over all regions r contained within z . By Euler's Formula, the number of all regions is equal to $q - p + 2$, so $|\mathcal{B}(G)| = b(G)$ and hence $\mathcal{B}(G)$ is a cycle basis.

The *interior* $int(P)$ of a nontrivial path P is the path minus its two endpoints. A *topological path* is any curve which is homeomorphic to the closed unit interval $[0, 1]$. A subset of the plane is *topologically path-connected* if any two of its points can be joined by a topological path which is the union of a finite number of straight line segments.

Lemma 5.1. *For any 2-connected plane graph G and $z \in Cyc(G)$, either $z = \partial r$ for some $r \in \mathcal{R}(G)$ or there exists $r \in \mathcal{R}(G)$ such that $z \cap \partial r$ is a nontrivial path.*

Proof. Assume $z \neq \partial r$ for all $r \in \mathcal{R}(G)$. Then $\partial r \cap z$ is a disjoint union of paths, some of which can consist of an isolated vertex. Let $\mathcal{R}_1, \mathcal{R}_2$ denote the sets of regions r in $\mathcal{R}(G)$ which are contained within z and for which $H = \partial r \cap z$ has at most 1, resp. at least 2, connected components. If \mathcal{R}_2 is empty, then for each edge e in z , the unique region r whose intersection with z includes e must satisfy $\partial r \cap z$ is a nontrivial path. Suppose \mathcal{R}_2 is nonempty. If P is any path subgraph of z joining two successive components of $z \cap \partial r$, then r separates the interior of P from the interior of $z - P$ in the sense that any topological path in the plane joining a point in $int(P)$ to a point in $int(z - P)$ must intersect ∂r . See [1, Lemma 4.1.2] for a formal proof. Let P^* be a path subgraph of z joining two successive components and of minimum possible length. Then for any edge e in P^* , there exists a unique region r' for which $e \in E(z \cap \partial r')$. If $z \cap \partial r'$ is not contained in P^* , then the path-connected set r' would allow a topological path between some point in the interior of P^* and some point in the interior of $z - P^*$. But this is impossible by our remark about the separating property of r . Hence, $z \cap \partial r' \subseteq P^*$. By minimality of P^* , r' must intersect z in a single component. Since the intersection contains an edge, it is a nontrivial path. \square

Theorem 5.2. *The basis $\mathcal{R}(G)$ of a 2-connected plane graph G is a robust basis; in fact, each cycle is the sum of a **wai** sequence with no repeated elements.*

Proof. Let t be the number of regions contained in z ; we prove the result by induction on t . For the basis case, $t = 1$ and the **wai** sequence has the single term z . If $t > 1$, then by

Lemma 5.1, there is a region r in $\mathcal{R}(G)$ contained in z such that $\partial r \cap z$ is a nontrivial path. Hence,

$$z' = z + \partial r \in Cyc(G)$$

and z' contains every region contained in z except for r . By the induction hypothesis, there is an ordering of the regions contained in z' so that

$$z' = \partial r_1 + \partial r_2 + \cdots + \partial r_{t-1}$$

and the sequence $(\partial r_1, \dots, \partial r_{t-1})$ is well-arranged w.r.t. intersection. Therefore, concatenating one more term, ∂r , gives the desired **wai** sequence of members of $\mathcal{R}(G)$. \square

6 Geodesic cycles can build all cycles

If z is any cycle (or path), let $\ell(z) := \text{length}(z)$ denote the number of edges. The *girth* $g = g(G)$ of a graph G is the length of a shortest cycle [1, p. 8]. A subgraph $H \subseteq G$ is *geodesic* (or *isometric*) if for all v, w vertices of H , $\text{dist}_H(v, w) = \text{dist}_G(v, w)$, where for $a, b \in V(H)$, $\text{dist}_H(a, b)$ is the distance from vertex a to vertex b in H - i.e., the length of a shortest a - b -path in H (and similarly for G). In K_4 , any 3-cycle is geodesic but any 4-cycle is not geodesic. It is clear that a cycle of minimum length is necessarily geodesic, so geodesic cycles always exist in any non-forest. The length of a geodesic cycle can't exceed twice the diameter by more than 1 as is shown in [1, proof of Prop. 1.3.2]. Hence, in a bipartite graph, every geodesic cycle has length at most twice the diameter.

Let $Geo(G)$ denote the set of geodesic cycles. We now show that for all graphs G , every cycle belongs to an iterated robust span of the geodesic cycles and bound the number of iterations needed; cf. [4].

Theorem 6.1. *For every graph G and every $z \in Cyc(G)$, one has*

$$z \in \rho^{\ell(z)-g(G)}(Geo(G)) \quad (6.1)$$

Proof. Suppose G is a fixed arbitrary graph and z is any cycle of G . We write $k = k(G, z) = \ell(z) - g(G)$. The statement of the theorem holds when $k = 0$ since then the cycle must be of minimum length, hence geodesic, and so in $\rho^0(Geo(G))$. Suppose the theorem holds when $k(G, z) < n$, where n is some positive integer and let z be a cycle of G with length $n + g(G)$ (i.e., so that $k(G, z) = n$). If z is not a geodesic cycle, there exist non-adjacent vertices v, w in z and a v - w -path P in G , intersecting z in exactly the two vertices v, w (and no edges) such that the length of P is less than the distance between v and w in z . Let P_1, P_2 be the two disjoint v - w -paths in z and put $z_i = P \cup P_i$, $i = 1, 2$. Since (z_1, z_2) is a **wai** sequence with sum z , $z \in \rho(\{z_1, z_2\})$. But both z_1 and z_2 are strictly shorter than z so by the inductive hypothesis, $z_1, z_2 \in \rho^{k'}(Geo(G))$, where

$$k' = \max\{\ell(z_1), \ell(z_2)\} - g(G) \leq k - 1.$$

Therefore, we have z in $\rho^{1+k'}(Geo(G)) \subseteq \rho^k(Geo(G))$ which completes the induction. \square

Let $c(G)$ denote the *circumference* of G , which is the length of a longest cycle.

Corollary 6.2. *For every graph G , $Cyc(G) \subseteq \rho^{c(G)-g(G)}(Geo(G))$.*

7 Characterizing geodesic cycles in hypercubes

The Cartesian product $G \times G'$ of graphs $G = (V, E)$ and $G' = (V', E')$ is the graph

$$V(G \times G') = V \times V'$$

with $[(v, v'), (w, w')] \in E(G \times G')$ iff $(v = w \ \& \ [v', w'] \in E(G'))$ or $(v' = w' \ \& \ [v, w] \in E(G))$ for all v, w in V and all v', w' in V' . The distance lemma (e.g., [10, p. 100]) generalizes this

$$d_{G \times G'}((v, v'), (w, w')) = d_G(v, w) + d_{G'}(v', w').$$

The hypercube Q_d is the d -fold Cartesian product of K_2 with itself; $V(Q_d) := \{0, 1\}^d$ consist of all binary d -tuples with $vw \in E(Q_d)$ exactly when v and w disagree in a unique coordinate. Thus, each edge is associated with a single active coordinate in $[d] := \{1, \dots, d\}$. We interpret the vertices as vectors or strings as convenient. As Q_d is bipartite with diameter d , the longest geodesic cycle has length not exceeding $2d$.

Write u_j for the vertex with a single 1 in the j -th coordinate so that

$$u_j : [d] \rightarrow \mathbf{Z}_2$$

is the characteristic function of the singleton subset $\{j\}$ of $[d]$.

Two vertices are *antipodal* (in the d -cube) if they have distance d (i.e., each coordinate is changed); two vertices are *diametrically opposite* (in an even-length cycle) if they are at maximum distance within the subgraph determined by the cycle. Two edges are *diametrically opposite* (in an even-length cycle) if they correspond to diametrically opposite vertices in the corresponding line graph. Let $\bar{0} := (0, 0, \dots, 0)$ be the 0-vector so $\bar{1} := (1, 1, \dots, 1)$ is its antipode. The vertices determine a group under coordinatewise modulo-2 addition. In fact, Q_d is the Cayley color graph of \mathbf{Z}_2^d w.r.t. $\{u_1, \dots, u_d\}$.

The graph enjoys a nice property:

Homogeneity If $e = vw$ is any edge of Q_d , there is an automorphism of Q_d carrying vw to $\bar{0}u_1$ and there is a color-preserving automorphism of Q_d carrying vw to $\bar{0}u_r$, where r is the active coordinate in the edge vw .

Recall that a walk (of length k) in a graph G is a sequence v_1, v_2, \dots, v_{k+1} of vertices of G such that, for each j , $1 \leq j \leq k$, v_j and v_{j+1} are adjacent in G . The walk is *closed* if $v_1 = v_{k+1}$. For any length- k v - w -walk $\omega = (v = v_1, v_2, \dots, v_{k+1} = w)$ in Q_d we define

$$\Phi(\omega) := (r_1, r_2, \dots, r_k) \in [d]^k,$$

where r_j is the active coordinate for the j -th edge $v_j v_{j+1}$ of the walk, $1 \leq j \leq k$. Conversely, given any vertex $v \in V(Q_d)$ and any k -tuple \mathbf{s} in $[d]^k$, $\mathbf{s} = (r_1, \dots, r_k)$, there is a unique length- k walk ω as above with $v_1 = v$ such that $\Phi(\omega) = \mathbf{s}$. Indeed, take $v_2 = v_1 + u_{r_1}$, $v_3 = v_2 + u_{r_2}$, etc. Also, $\omega^{op} := (v_{k+1}, v_k, \dots, v_1)$ is a w - v -path and $\Phi(\omega^{op})$ is the reversal of $\Phi(\omega)$. For each vertex v in the hypercube, there is a one-to-one length-preserving correspondence between walks starting at v and sequences in $[d]^k$.

By the distance lemma, a walk ω , in Q_d , corresponds to a geodesic path if and only if $\Phi(\omega)$ is a permutation of some nonempty subset of $[d]$. Hence, one can characterize walks in the hypercube which correspond to geodesic cycles. Let $a * b$ denote concatenation of strings a, b .

Theorem 7.1. *A walk ω in Q_d is a geodesic cycle if and only if $\exists S \subseteq [d]$, $|S| \geq 2$, such that*

$$\Phi(\omega) = \sigma * \sigma, \quad (7.1)$$

where σ is a string which is a permutation of the elements of S .

Proof. Let z be any geodesic cycle in Q_d . Then for any two diametrically opposite vertices v, w of z , the two distinct v - w -paths P, P' contained in z must both be geodesic paths. Let τ be the closed walk P' followed by P^{op} . By the remark above, both P and P' correspond to permutations of sets S, S' , resp. But $S = S'$ as the walk is closed. Hence, any r in $[d]$ appears either twice or not at all in the string $\Phi(\tau)$. If the two appearances of some $r \in [d]$ were not at diametrically opposite edges e, e' of z , then there exist diametrically opposite vertices $v', w' \in V(z)$ such that both e and e' are contained in one of the two geodesic paths in z which is impossible.

Conversely, for a walk of length $2j$ satisfying the permutation condition, any subpath of length $\leq j$ corresponds to a permutation and is geodesic so the cycle is also geodesic. \square

8 Building geodesic cycles in the hypercube

Since Q_d is bipartite, shortest cycles have length ≥ 4 . For distinct i, j in $[d]$, the closed walk $\omega = (v, v + u_i, v + u_i + u_j, v + u_j, v)$ is a 4-cycle and $\Phi(\omega) = (i, j, i, j)$. Moreover, any 4-cycle, being geodesic, must be of this form. Call 4-cycles *squares* and let $Sq(Q_d)$ be the set of all squares in Q_d . For convenience, during this section we shall write ij for any square s such that $\Phi(s) = (i, j, i, j)$, for distinct $i, j \in [d]$.

We will use the notion of a *strip of squares* in some graph G . This is a sequence

$$\sigma = (s_1, s_2, \dots, s_k), \quad (8.1)$$

where each s_j is a distinct 4-cycle subgraph of G , no two squares intersect unless one is the successor of the other, and for $2 \leq j \leq k - 1$, s_j intersects its predecessor s_{j-1} and its successor s_{j+1} in disjoint, that is, *opposite*, edges. The *length* of a strip of squares is the number of squares in the sequence. If the length of a strip of squares σ as in (8.1) is $k \geq 2$, then we call the strip *nontrivial*.

For a nontrivial strip there are distinguished *start* and *stop* edges for σ : the start edge is the edge of s_1 opposite to the edge $s_1 \cap s_2$, and the stop edge is the edge of s_k opposite to the edge $s_{k-1} \cap s_k$. The *ties* of a nontrivial strip of squares is the set consisting of the start and stop edges, together with all *intersection edges* $s_j \cap s_{j+1}$, $1 \leq j \leq k - 1$. Let $G(\sigma) := \bigcup_{j=1}^k s_j$ denote the *graph determined by the strip of squares*. A strip of length 1 has graph which is a 4-cycle and so is isomorphic to $Q_1 \times P_2$. The following results can be easily proved by induction on k .

Lemma 8.1. *Let σ be a strip of squares of length $k \geq 2$. Then (i) $G(\sigma)$ is isomorphic to $Q_1 \times P_{k+1}$, (ii) the ties of σ correspond to the edges $Q_1 \times v$ for $v \in V(P_{k+1})$, and (iii) the squares s_1, \dots, s_k are a robust basis of $G(\sigma)$.*

Lemma 8.2. *In a hypercube, all ties of a strip of squares have the same active coordinate.*

In the following, assume that $d \geq 3$, else everything is trivial. For $1 \leq j \leq d$, define strips σ_1^j in Q_d as follows. Let s_2 be the unique 12 square which contains $\bar{0}$, put $\sigma_1^1 = s_2$, and let $[u_2, u_2 + u_1]$ be the “stop edge” of s_2 . For $2 \leq j \leq d$, having defined σ_1^{j-1} let

$$\sigma_1^j := \sigma_1^{j-1} * s_j,$$

where s_j is the unique $1j$ square through the vertex $u_2 + \dots + u_{j-1}$. Let z_1^j be the sum of the squares in σ_1^j . that is,

$$z_1^j := s_2 + \dots + s_j.$$

We prove by induction on j that

$$\Phi(z_1^j) = (1, \dots, j, 1, j, j - 1, \dots, 2). \tag{8.2}$$

For $j = 2$ the result holds and for $2 \leq j \leq d - 1$, adding s_{j+1} to z_1^j replaces the stop edge of σ_1^j with a sequence of three edges active consecutively in coordinates $j + 1, 1, j + 1$. Thus,

$$\Phi(z_1^{j+1}) = \Phi(z_1^j + s_{j+1}) = (1, \dots, j + 1, 1, j + 1, j, j - 1, \dots, 2),$$

establishing the induction. Let $\sigma_1 = \sigma_1^d$ and put $z_1 := z_1^d$. Then

$$\Phi(z_1) = (1, 2, \dots, d, 1, d, d - 1, \dots, 3, 2).$$

Thus, z_1 is the union of two paths P, P_1 from $\bar{0}$ to $\bar{1}$, where $\Phi(P) = (1, 2, \dots, d)$ and $\Phi(P_1) = (2, \dots, d, 1)$. Hence, z_1 is a cycle of length $2d$, but it is not geodesic since the intersection edges of σ_1 join non-adjacent vertices of z_1 .

We now show that every geodesic cycle in Q_d is in the robust span of the set of squares by continuing the above construction.

Theorem 8.3. *Let z be any geodesic cycle of length $2j$ in Q_d , $d \geq 3$. Then there exists a wai z -sequence of distinct squares z_1, z_2, \dots, z_m , where $m = j(j - 1)/2$. So $z \in \rho(Sq(Q_d))$.*

Proof. By the homogeneity of Q_d , it suffices to show that, for each positive integer $d \geq 3$, the unique geodesic cycle z in Q_d through $\bar{0}$ with $\Phi(z) = (1, 2, \dots, d, 1, 2, \dots, d)$ is the sum of a sequence σ of squares which is well-arranged w.r.t. intersection.

For $1 \leq j \leq d - 1$, let σ_j denote a sequence of squares of length $d - j$ of the form

$$\sigma_j := (j \ j + 1, j \ j + 2, \dots, j \ d);$$

put $\sigma := \sigma_1 * \sigma_2 * \dots * \sigma_{d-1}$, the concatenation. Each σ_j is a strip of squares, the strips are disjoint in the sense that no square belongs to more than one, and every square does belong to a strip. We will show that this sequence of squares is wai and has sum z . In the sequel, as we add squares to $z_1 = P \cup P_1$, the path P will be unchanged while the path P_1 will be successively transformed, but without changing its length.

The sequence $\sigma_2 = (s_{23}, s_{24}, \dots, s_{2d})$ is defined as follows. Let s_{23} be the unique 23 square which includes the vertex $\bar{0}$; note that s_{23} intersects P_1 in the path

$$p_{1,23} = (\bar{0}, u_2, u_2 + u_3)$$

which are the first two edges of P_1 . Adding s_{23} to z_1 produces a new cycle $z_{1,23}$ which is the union of P and the path $P_{1,23}$ obtained by replacing $(0, u_2, u_2 + u_3)$ by $(0, u_3, u_2 + u_3)$ in P_1 . Now there is a unique 2^4 square s_{24} which includes the vertex u_3 , and s_{24} intersects $P_{1,23}$ in the path

$$p_{1,24} = (u_3, u_2 + u_3, u_2 + u_3 + u_4)$$

which constitutes the second and third edges of $P_{1,23}$. Put $z_{1,24} = z_{1,23} + s_{24}$. Then $z_{1,24} = P \cup P_{1,24}$, where

$$P_{1,24} = (\bar{0}, u_3, u_3 + u_4, \sum_{j=2}^4 u_j, \sum_{j=2}^5 u_j, \dots, \sum_{j=2}^d u_j, \sum_{j=1}^d u_j).$$

Now there is a unique 2^5 square through the vertex $u_3 + u_4$ intersecting $P_{1,24}$ in its third and fourth edges, and so forth up to $2d$. Adding the sum z'_2 of the squares in σ_2 to z_1 produces

$$\Phi(z_2) = \Phi(z_1 + z'_2) = (1, 2, \dots, d, 1, 2, d, d - 1, \dots, 3).$$

Proceeding in this fashion, $z = z_d$, $\Phi(z) = (1, \dots, d, 1, \dots, d)$, and σ is a **wai** z -sequence as each square after those in the initial subsequence σ_1 intersects the sum of the preceding squares in a path of length 2. \square

Thus, the set $Sq(Q_d)$ of all squares in Q_d robustly spans the set $Geo(Q_d)$ of all geodesic cycles in Q_d ,

$$Geo(Q_d) \subseteq \rho(Sq(Q_d)).$$

But $Sq(Q_d)$ contains $d(d-1)2^{d-3}$ squares, which is asymptotically $d/4$ times the size of a cycle basis. Indeed, $b(Q_d) = 1 + d2^{d-1} - 2^d = 1 + (d-2)2^{d-1}$; e.g., [7, pp 22–23, 38–39].

In the graph $Q_1 \times G$, as $V(Q_1) = \{0, 1\}$, G is isomorphic to $0 \times G$ under the mapping

$$v \mapsto (0, v),$$

and we write $0 \times H$ for the image under this isomorphism of a subgraph H of G in $0 \times G$ and similarly for $1 \times H$. Also, if $e \in E(G)$, we write e for the corresponding subgraph (which is isomorphic to Q_1), and for $e \in E(G)$, each subgraph $Q_1 \times e$ in $Q_1 \times G$ is a 4-cycle.

We shall use the decomposition

$$E(Q_d) = E(Q_d^0) \cup E(Q_d^1) \cup E_d, \tag{8.3}$$

where

$$Q_d^0 = 0 \times Q_{d-1} \quad \text{and} \quad Q_d^1 = 1 \times Q_{d-1}$$

denote the “bottom” and “top” subgraphs of the d -cube,

$$E_d := E(Q_d) \setminus (E(Q_d^0) \cup E(Q_d^1))$$

denotes the “side” edges of the d -cube. Similarly, there is a decomposition of the squares

$$Sq(Q_d) = S_0 \cup S_1 \cup Sq'_d, \tag{8.4}$$

where $\mathcal{S}_j := Sq(Q_d^j)$ for $j = 0, 1$ and $Sq'_d = Sq(Q_d) \setminus (\mathcal{S}_0 \cup \mathcal{S}_1)$ (the set of all “side” squares of Q_d). Each element s in Sq'_d is the product of Q_1 with the unique edge e in $E(Q_{d-1})$ corresponding to the intersection of s with Q_d^0 .

Now we can define the *Kainen basis* of Q_d [12]. Let $\mathcal{B}_K(Q_d)$ be the following recursively given collection of squares in Q_d . For $d = 0$ and $d = 1$, the set is empty, and $\mathcal{B}_K(Q_2) = \{Q_2\}$. Having defined $\mathcal{B}_K(Q_{d-1})$ for $d \geq 3$, put

$$\mathcal{B}_K(Q_d) := \mathcal{B}_K(Q_d^0) \cup Sq'_d, \tag{8.5}$$

where $\mathcal{B}_K(Q_d^0)$ is the cycle basis of Q_d^0 corresponding to $\mathcal{B}_K(Q_{d-1})$ under the natural isomorphism. For example, $\mathcal{B}_K(Q_4)$ consists of five squares in one of the 3-cube faces and all twelve side-squares.

Lemma 8.4. *The set $\mathcal{B}_K(Q_d)$ is a cycle basis for Q_d for $d \geq 2$.*

Proof. By induction on d ; trivial for $d = 2$. Now let $d \geq 3$. Then $\mathcal{B}_K(Q_d)$ is independent since by induction the squares in $\mathcal{B}_K(Q_d^0)$ in the decomposition (8.5) are independent, while the squares in the sides, Sq'_d , are independent of those in the bottom and also of each other as each contains a unique edge in Q_{d-1} . But $\mathcal{B}_K(Q_d)$ has the cardinality of a basis (and so is a basis). Indeed, using (8.5), we have

$$|\mathcal{B}_K(Q_d)| = |\mathcal{B}_K(Q_{d-1})| + |E(Q_{d-1})| = 1 + (d-3)2^{d-2} + (d-1)2^{d-2} = 1 + (d-2)2^{d-1},$$

where the second equality again uses the inductive assumption that $\mathcal{B}_K(Q_{d-1})$ is a basis. \square

We now show that this basis is sufficient to robustly span every square in Q_d .

Theorem 8.5. *For $d \geq 2$, $Sq(Q_d) \subseteq \rho(\mathcal{B}_K(Q_d))$.*

Proof. As before, the basis case $d = 2$ is trivial. Every square in Q_d which is on the bottom is in $\rho(\mathcal{B}_K(Q_{d-1}))$ by the induction hypothesis and $\rho(\mathcal{B}_K(Q_{d-1})) \subseteq \rho(\mathcal{B}_K(Q_d))$ and the side-squares are in $\mathcal{B}_K(Q_d)$ by definition. Thus, we only need to take care of the squares s on the top; that is, $s \in \mathcal{S}_1 = Sq(Q_d^1)$. Each such square is contained in a unique subgraph Q_3 of Q_d such that $s = Q_3^1$. Let $s' = Q_3^0$. Then $s' \in Sq(Q_d^0)$ so by induction there exists σ' a **wai** s' -sequence in $\mathcal{B}_K(Q_{d-1}) \subseteq \mathcal{B}_K(Q_d)$, and we may append to this sequence the 4 side-squares of Q_3 (in any order) to produce a **wai** sequence in $\mathcal{B}_K(Q_d)$ which sums to s . \square

Using Theorem 6.1, we see that for hypercubes every cycle is in the iterated robust closure of $\mathcal{B}_K(Q_d)$.

Corollary 8.6. *For $d \geq 2$ and z any cycle in Q_d , $z \in \rho^{\ell(z)-2}(\mathcal{B}_K(Q_d))$.*

9 Application to commutativity in groupoids

The original motivation for building up cycles using **wai** sequences was to show that *the commutativity of certain diagrams in a groupoid category can be inferred from the commutativity of a small minority of the faces of the diagram* [13]. After defining these terms, we use the results of the previous sections of this paper to prove the above claim and we show that a similar inference *cannot* be made using only **wat** sequences.

For a full discussion of the concept of a category, see, e.g., Mac Lane [17]. Briefly, a (small) category \mathcal{C} consists of a set $Obj(\mathcal{C})$ of objects, a set $Mor(\mathcal{C})$ of morphisms, and two functions (“domain” and “co-domain”)

$$dom, cod : Mor(\mathcal{C}) \rightarrow Obj(\mathcal{C}).$$

(A morphism thus models the notion of a structure-preserving function between two mathematical objects.) An ordered pair (a, b) of morphisms is called *composable* if $dom(b) = cod(a)$. In addition, \mathcal{C} is required to have a *law of composition* which is a function from the set of composable morphism pairs to the set of morphisms denoted by

$$(a, b) \mapsto ab,$$

where $dom(ab) = dom(a)$ and $cod(ab) = cod(b)$. Often the notation for composition is in reverse order, emulating composition of functions.

It is convenient to write $\mathcal{C}(x, y)$ for the set of all morphisms in \mathcal{C} with domain x and co-domain y , and $a : x \rightarrow y$ is another notation for $a \in \mathcal{C}(x, y)$. Typically, the set of morphisms between a given pair of objects has many elements but this set can also be empty or a singleton. For any triple x, y, z of objects of \mathcal{C} the law of composition gives a function $\mathcal{C}(x, y) \times \mathcal{C}(y, z) \rightarrow \mathcal{C}(x, z)$.

Composition in a category \mathcal{C} must be *associative*: if (a, b) and (b, c) are composable pairs of morphisms in \mathcal{C} , then (ab, c) and (a, bc) are both composable pairs and we require $(ab)c = a(bc)$. It follows by induction that one can define a unique composition $a_1 a_2 \cdots a_k$ for any finite sequence (a_1, a_2, \dots, a_k) of morphisms for which each successive pair (a_i, a_{i+1}) is composable. The final condition in the definition of a category \mathcal{C} is that for each object x , there is an *identity morphism* $1_x \in \mathcal{C}(x, x)$ such that if x, y are any two objects in \mathcal{C} , then for any morphism a in $\mathcal{C}(x, y)$,

$$a1_y = a = 1_x a.$$

Typical examples of a category are a set of topological spaces and continuous maps, or a set of vector spaces and linear maps, etc., provided that the composition of any two morphisms remains in the set, and of course that the identity morphisms are included.

A directed multigraph $D = (V, A, \Psi)$ is a set V of vertices, a set A of arcs, and a mapping $\Psi : A \rightarrow V \times V$ which associates to each arc an ordered pair (v, w) , where v is its source and w its target. Arc-pairs (a, a') are *composable* if the target of a is the source of a' and an arc sequence is a *dipath* if each successive pair is composable. A v - w -dipath is a dipath where the first arc has source v and the last arc has target w . A *diagram* δ in a category \mathcal{C} is a directed multigraph $D = (V, A, \Psi)$ (called the *scheme*) and functions f_V, f_A such that

$$f_V : V \rightarrow obj(\mathcal{C}), \quad f_A : A \rightarrow mor(\mathcal{C})$$

such that if $\Psi(a) = (v, w)$ and $f_V(v) = x, f_V(w) = y$, then $f_A(a) \in \mathcal{C}(x, y)$. Unless necessary, we shall not distinguish between vertices of the diagram and objects of the category, nor between arcs and morphisms. A dipath in the diagram corresponds to a composable sequence of morphisms in the category.

A *face* of the diagram is a distinct pair of v - w -dipaths, for some objects v, w , and this face is *commutative* if the two paths give rise to the same v - w -morphism, using the law of composition in the category. The face is said to “commute.” The diagram is commutative if all of its faces commute.

The importance of commutative diagrams is that they can be used to define certain algebraic properties, such as co-associativity, and a basic theory of algebraic syntax can be built up from commutative diagrams - see, e.g., [18] and [5]. For an explicit recent example from computer science, see [21, p. 13, diagram for Remark 1.2.6].

A *groupoid* is a category \mathcal{G} in which every morphism is invertible; i.e., for every morphism a , there is a unique morphism a^{-1} such that composition in either order gives 1. Groupoids are generalizations of ordinary groups and they have become of some interest in a number of fields recently, spreading far from initial applications in topology; see, e.g., [23].

A diagram in a groupoid \mathcal{G} can be extended by including for every arc a in the diagram a reverse arc a^{op} and if the morphism α is assigned to a then in the extended diagram, the inverse morphism α^{-1} is assigned to a^{op} . If the original diagram commutes, then so does the extended diagram and the reverse implication is trivial. Hence, we shall assume that all diagrams in a groupoid category are symmetric.

Any undirected cycle in a groupoid diagram can be parsed into a closed walk by choosing one of the vertices and then proceeding either clockwise or counterclockwise. The cycle is said to *g-commute* if the composition along the walk is the identity; this does not depend on how the cycle is parsed; e.g., if $abc = 1$, then $1 = a^{-1}1a = bca$, etc. A diagram in a groupoid category will be called *g-commutative* if and only if all of its cycles g-commute.

Lemma 9.1. *If a groupoid diagram is g-commutative, then it is commutative.*

Proof. Suppose we have a g-commutative diagram. Let P, Q be two directed paths from v to w in the diagram and let α, β be the morphisms induced by P and Q , resp. It suffices to assume P and Q intersect only in v and w . Let z be the cycle formed by the union of P and Q , parsed to begin at v , follow P , and then Q (in reverse order). Then z induces $\alpha\beta^{-1}$. So the morphism induced by z is 1 if and only if $\alpha = \beta$. Thus, g-commutativity implies ordinary commutativity. □

The converse is false. For example, one can have four objects w, x, y, z with morphisms

$$f : w \rightarrow x, g : y \rightarrow x, h : y \rightarrow z, i : w \rightarrow z.$$

The corresponding diagram is automatically commutative because it has no faces but if it is in a groupoid it is g-commutative only if $fg^{-1}hi^{-1} = 1$, i.e., $fg^{-1} = ih^{-1}$.

Proposition 9.2. *The g-commutative cycles in a groupoid diagram form a cooperative set.*

Proof. Let z, z' be commutative cycles in a groupoid diagram, with $P = z \cap z'$ a nontrivial path from v to w . Let α be composition along P , β along $z - P$, and β' along $z' - P$. Then

$$\alpha\beta = 1_v \quad \text{and} \quad \alpha\beta' = 1_v$$

But also $\beta^{-1}\alpha^{-1} = 1_v$. Hence,

$$1_v = \beta^{-1}\alpha^{-1}\alpha\beta' = \beta^{-1}\beta'.$$

Thus, $z + z'$, which is the cycle $(z - P) \cup (z' - P)$ is commutative. □

Using Lemma 4.1 and Proposition 9.2, by Theorem 6.1 and Corollary 8.6, we get:

Corollary 9.3. *A diagram in a groupoid is g -commutative if its geodesic cycles g -commute. A Q_d -diagram in a groupoid is g -commutative if each $z \in \mathcal{B}_K(Q_d)$ g -commutes.*

The sum of two commutative cycles can be a noncommutative cycle when the cycles intersect in more than one nontrivial path. Label the vertices of a complete graph on 4 vertices A, B, C, D in clockwise order, where all vertices correspond to a single object X in the groupoid, and take all morphisms to be 1_X , except for the two diagonal arcs from C to A and from D to B , respectively, which represent some morphism $x : X \rightarrow X$, where $x^2 \neq 1_X$. (This is a mild condition on the groupoid, satisfied for instance by the symmetric group S_n for $n \geq 3$.) Consider the two squares A, B, C, D and A, B, D, C ; going around the first square gives a 4-fold composition of identities, while the second square gives

$$x \circ 1_X \circ x^{-1} \circ 1_X = 1_X.$$

The two squares intersect in edges AB and CD and the sum of the two squares is the square A, D, B, C which induces the morphism $x^2 \neq 1_X$. Thus, propagation of commutativity to the sum can fail when cycles intersect in two disjoint paths.

10 Remarks

This section is essentially an appendix which compares other types of cycle basis to the one we have constructed. We thank one of the referees for suggesting some of the references given here and for other helpful comments.

A general method for generating cycle bases for Cartesian product graphs, was described by Hammack [6]. It constructs a cycle basis for $G \times G'$ for each pair of graphs G, G' with given spanning trees T, T' and cycle bases $\mathcal{B}, \mathcal{B}'$ for G, G' , resp. Then there is a *Hammack basis*

$$\mathcal{B}_H(G, \mathcal{B}, T, G', \mathcal{B}', T') := \mathcal{F} \cup \mathcal{G} \cup \mathcal{G}', \text{ where}$$

$$\mathcal{F} := \{e \times e' : e \in E(T), e' \in E(T')\},$$

$$\mathcal{G} := \{z \times y : z \in \mathcal{B}, y \in V(G')\},$$

$$\mathcal{G}' := \{x \times z' : x \in V(G), z' \in \mathcal{B}'\}.$$

For the hypercube Q_d , it is reasonable to take $G = Q_{d-1}$ and $G' = Q_1$. So $T' = Q_1$ but T could be an arbitrary spanning tree of Q_{d-1} . There is a natural recursive choice for T a spanning tree of Q_{d-1} when $d \geq 3$: just keep adding in all of the “side” edges. For example, the spanning tree constructed for Q_3 looks like a “U” standing on 4 legs. If $d \geq 4$, the Hammack bases (allowing all possible spanning trees) do not include the Kainen basis since the latter uses only bottom squares, but all $(d-1)2^{d-2}$ side-squares, while the former type use both top and bottom squares but only $2^{d-1} - 1$ side squares. For example, $\mathcal{B}_H(Q_4)$ will have 5 squares in both Q_4^0 and Q_4^1 but only 7 side-squares (cf. supra Lemma 8.4). Other iterated Cartesian product graphs have been explicitly studied, e.g., in [9].

Minimum cycle bases (in the sense of having minimum total sum of the lengths of their members) for Cartesian product graphs were constructed by Imrich and Stadler [11]. Like the Hammack bases, these depend on a spanning tree and a cycle basis, and also a vertex, for each factor of a Cartesian product. The *Imrich-Stadler basis* for $G \times G'$ has the form

$$\mathcal{B}_{IS}(G, \mathcal{B}, T, x, G', \mathcal{B}', T', y) := \mathcal{B}_{\square} \cup \mathcal{G}^y \cup \mathcal{G}'_x,$$

where $\mathcal{B}, T, x (\mathcal{B}', T', y)$ are cycle basis, spanning tree, and vertex for G (and for G' , resp.), and

$$\begin{aligned} \mathcal{B}_\square &:= \{e \times e' : e \in E(T), e' \in E(G')\} \cup \{e \times e' : e \in E(G), e' \in E(T')\}, \\ \mathcal{G}^y &:= \{z \times y : z \in \mathcal{B}\}, \\ \mathcal{G}'_x &:= \{x \times z' : z' \in \mathcal{B}'\}. \end{aligned}$$

If $G' = T'$, then $\mathcal{B}_\square = \{e \times e' : e \in E(G), e' \in E(T')\}$ and \mathcal{G}'_x is empty. So $\mathcal{B}_{IS}(G \times T')$ doesn't depend on T or x . When $G = Q_{d-1}$ and $T = Q_1$, this is just $\mathcal{B}_K(Q_d)$ so \mathcal{B}_{IS} is a generalization of \mathcal{B}_K . When the cycle bases of G and G' are minimal and both G and G' have girth at least 4, [11, Thm. 8] shows that $\mathcal{B}_{IS}(G \times G')$ is minimal as well. When G' is a tree, we can get a robustness-related result. In the graph $G \times T$, T a tree, a cycle z contained in a subgraph $G \times v'$ for some $v' \in V(T)$ is called a *level cycle*.

Proposition 10.1. *Let G be a graph with spanning tree T , cycle-basis \mathcal{B} , and $v_0 \in V(G)$. If \mathcal{B} is robust, then every level-cycle in $T \times G$ is in the robust span of the basis \mathcal{B}'*

$$\mathcal{B}' := \{v_0 \times z : z \in \mathcal{B}\} \cup \{e' \times e : e' \in E(T), e \in E(G)\}.$$

Proof. Suppose that T has t vertices and G has p vertices and q edges. Then it is easy to calculate that $b(T \times G) = qt - p + 1 = |\mathcal{B}'|$ and \mathcal{B}' is a basis by [11, Thm. 4].

We prove the robustness claim by induction on the number t of vertices of T . When $t = 1$, $\mathcal{B}' = \mathcal{B}$ so the result holds. Let T be any tree with t vertices, $t > 1$. Then there is a vertex v in T such that v has degree 1. Let $T' = T - v$. Now consider any level cycle z' in $T \times G$. If z' is in $w \times G$ for some $w \in V(T')$, then by the induction hypothesis, there is a z' -sequence from \mathcal{B}' which is **wai**. Otherwise, if z' is contained in $v \times G$, then $z' = v \times z$ for some $z \in Cyc(G)$. Let v' be the unique vertex of T which is adjacent to v . For the cycle $z'' = v' \times z$, there is a z'' -sequence σ' from \mathcal{B}' which is **wai**. If the edges in z are followed in a closed walk, say e_1, \dots, e_k and if $e = vv'$, then

$$\sigma = \sigma' * (e \times e_1, \dots, e \times e_k)$$

is a **wai** sequence from \mathcal{B}' which sums to z' . □

Recently, convex cycle bases were studied in [8]. A subgraph H of G is *convex* if for every minimal-length G -path P between vertices $v, w \in V(H)$, P is a subgraph of H . In a hypercube, squares are convex subgraphs. The Imrich-Stadler basis for $G \times H$ is convex when the bases for G and H are convex. They note that the only convex cycles in a hypercube are the squares. Geodesic cycles in the hypercubes are a special case of *planar partial cube* and the latter have been characterized in [20].

References

- [1] R. Diestel, *Graph Theory*, Springer, Berlin, 2005.
- [2] E. T. Dixon and S. E. Goodman, An algorithm for the longest cycle problem, *Networks* **6**(2) (1976), 139–149.
- [3] U. Doğrusöz and N. Krishnamoorthy, Enumerating all cycles of a planar graph, *Parallel Algorithms and Applications* **10** (1996), 21–36.

- [4] A. Georgakopoulos and P. Sprüssel, Geodesic topological cycles in locally finite graphs, *The Electronic Journal of Combinatorics* **16** (2009), #R144 (18 pages).
- [5] J. A. Goguen, A categorical manifesto, in *Mathematical Structures in Computer Science* **1** (1991), 49–67; <http://www.cs.ox.ac.uk/files/3395/PRG72.pdf>
- [6] R. Hammack, Cyclicity of graphs, *Journal of Graph Theory* **32**(2) (1999), 160–170.
- [7] F. Harary, F., *Graph Theory*, Addison-Wesley, Reading, MA, 1970.
- [8] M. Hellmuth, J. Leydold and P. F. Stadler, Convex cycle bases, *Ars Math. Contem.* **7**(2014), 123–140.
- [9] W. Imrich and S. Klavžar, *Product Graphs: Structure and Recognition*, Wiley, New York, 2000.
- [10] W. Imrich, S. Klavžar, and D. F. Rall, *Graphs and Their Cartesian Product*, A. K. Peters, Wellesley, MA, 2008.
- [11] W. Imrich and P. F. Stadler, Minimal cycle bases of product graphs, *Australasian J. of Combinatorics* **26** (2002) 233–244.
- [12] P. C. Kainen, On robust cycle bases, Proc. 9th Quadr. Conf. in Graph Theory, June 2000, Y. Alavi et al., Ed., in *Electr. Notes in Discr. Math.* **11** (2002), 430–437 (art. 38, pp. 1–8).
- [13] P. C. Kainen, Isolated squares in hypercubes and robustness of commutativity, *Cahiers de Topologie et Géométrie Différentielle Catégoriques* **XLIII** (2002), 213–220.
- [14] P. C. Kainen, Graph cycles and diagram commutativity, pp. 177–238, in *Diagrammes*, Suppl. aux vol. **67+68**, Paris, 2012.
- [15] K. Klemm and P. F. Stadler, Statistics of cycles in large networks, *Phys. Rev. E* **73** (2006) 025101.
- [16] K. Klemm and P. F. Stadler, A note on fundamental, non-fundamental, and robust cycle bases, *Discrete Applied Mathematics* **157** (2009), 2432–2438.
- [17] S. Mac Lane, *Categories for the working mathematician*, Springer, New York, 1971
- [18] F. Orejas, On the representation of data types, in *Formalization of Programming Concepts*, Lecture Notes in Computer Science, Vol. 107, 1981, 419–431.
- [19] P.-J. Ostermeier, M. Hellmuth, K. Klemm, J. Leydold and P. F. Stadler, A note on quasi-robust cycle bases, *Ars Mathematica Contemporanea* **2** (2009), 231–240.
- [20] I. Peterin, A characterization of planar partial cubes, *Discrete Math.* **308**(2008), 6596–6609.
- [21] B. C. Pierce, *Basic Category Theory for Computer Scientists*, MIT Press, 1991.
- [22] M. M. Sysło, On cycle bases of a graph, *Networks* **9** (1979), 123–132.
- [23] A. Weinstein, Groupoids: Unifying Internal and External Symmetry, *Notices Amer. Math. Soc.* **43** (1996), 744–752.
- [24] A. White, *Graphs of Groups on Surfaces*, Mathematics Studies 188, North-Holland, Amsterdam, 2001.