

Book embeddings of Reeb graphs

Vitaliy Kurlin

<http://kurlin.org>

vitaliy.kurlin@durham.ac.uk

Department of Mathematical Sciences
Durham University, Durham DH1 3LE, UK

December 5, 2013

Abstract

Let X be a simplicial complex with a piecewise linear function $f : X \rightarrow \mathbb{R}$. The *Reeb graph* $\text{Reeb}(f, X)$ is the quotient of X , where we collapse each connected component of $f^{-1}(t)$ to a single point. Let the nodes of $\text{Reeb}(f, X)$ be all homologically critical points where any homology of the corresponding component of the level set $f^{-1}(t)$ changes. Then we can label every arc of $\text{Reeb}(f, X)$ with the Betti numbers $(\beta_1, \beta_2, \dots, \beta_d)$ of the corresponding d -dimensional component of a level set. The homology labels give more information about the original complex X than the classical Reeb graph. We describe a canonical embedding of a Reeb graph into a multi-page book (a star cross a line) and give a unique linear code of this book embedding.

1 Our contributions and related work

Reeb graphs are used in image analysis [1] as simple representations of shapes. The Reeb graph is an abstract graph that describes the linking structure of connected components in level sets of a real-valued function on a complex. Briefly, a complex is a union of vertices, edges, triangles, tetrahedra and so on. Formal Definition 2.5 of a Reeb graph in section 2 doesn't specify any canonical way to visualize this abstract graph in a low dimensional space.

Problem 1.1. *Is it possible to canonically embed (draw without any self-intersections) and encode any general Reeb graph in a low-dimensional space?*

Here we briefly introduce auxiliary concepts, see all definitions in sections 2-3. A cycle C in a graph is independent of other cycles C_1, \dots, C_k if

(briefly) $C \subset \cup_{i=1}^k C_i$ and every edge of $(\cup_{i=1}^k C_i) - C$ is covered twice. For instance, the theta-graph θ has two independent cycles (small loops) whose union covers the large cycle once and the middle arc twice. In general, m multiple arcs between the same nodes generate $m - 1$ independent cycles.

Embedding Theorem 1.2 requires a Reeb graph simplified by all steps:

- (1) forget about every node of degree 2 and then merge its two arcs;
- (2) replace all multiple arcs between the same nodes by a single arc;
- (3) remove any subtree (a subgraph without cycles), but keep its root.

Let T_k be the star graph with k rays. Then $T_k \times \mathbb{R}$ is a book with k pages.

Theorem 1.2. *Let a Reeb graph have at most m independent cycles with a common arc. If a simplified Reeb graph has n nodes, then the original Reeb graph can be canonically embedded into the book $T_{(m+1)\max\{1, n-2\}} \times \mathbb{R} \subset \mathbb{R}^3$.*

The embedding in Theorem 1.2 is canonical in the sense that different embeddings of the same graph can be continuously deformed to each other. In a book embedding of a Reeb graph, all nodes lie in the binding axis of the book and each arc lies in a single page. A book embedding can be linearly encoded, so we may work with easy codes instead of abstract Reeb graphs.

Almost all results about book embedding focus on undirected graphs when nodes can lie in any order in the binding axis. There is a linear time algorithm to embed any planar graph in 4 pages [12]. A general undirected graph with n nodes can be embedded in $\lceil \frac{n}{2} \rceil$ pages [3]. We could find results on book embeddings (stack layouts) of general directed acyclic graphs only for a small number of pages [6]. The full version of Theorem 1.2 in section 4 extends [6, Theorem 2.3], where any graph with only one cycle is embedded into 2 pages. In section 5 we review other work on computing Reeb graphs.

Our methods for book embeddings were inspired by basic embeddings of graphs [8] and by a reduction of the topological classification for graphs embedded in \mathbb{R}^3 to word problems in finitely presented semigroups [9].

2 Simplicial complexes and their Reeb graphs

Definition 2.1. A *simplicial complex* is a finite set V of vertices and a collection of subsets $\sigma \subset V$ called *simplices* such that all subsets of a simplex are also simplices. The dimension of a simplex $\sigma = \{v_1, \dots, v_k\}$ is $k - 1$. The dimension of a complex is the highest dimension of all its simplices. Any simplex inherits the Euclidean topology from this geometric realization:

$$\Delta^d = \{(t_0, \dots, t_d) \in \mathbb{R}^{d+1} \mid t_0 + t_1 + \dots + t_d = 1 \text{ and all } t_i \geq 0\}.$$

Then we can define the topology on any simplicial complex by gluing all its simplices along their common subsimplices, so we glue σ, τ along $\sigma \cap \tau$.

Definition 2.2. The *Euler characteristic* of a simplicial complex X is the alternating sum $\chi(X) = \sum_{i \geq 0} c_i(X)$, where $c_i(X)$ is the number of i -dimensional simplices of X . The sum is finite as X consists of finitely many simplices.

Definition 2.3. Two maps $f, g : X \rightarrow Y$ are *homotopic* if there is a continuous map $\Phi : X \times [0, 1] \rightarrow Y$, such that $\Phi(x, 0) = f(x)$ and $\Phi(x, 1) = g(x)$ for any $x \in X$, so f, g are connected by a continuous family of maps $\Phi(\cdot, t) : X \rightarrow Y, t \in [0, 1]$. Two simplicial complexes X, Y have the same *homotopy type* (are *homotopy equivalent*) if there are continuous functions $f : X \rightarrow Y$ and $g : Y \rightarrow X$ such that $g \circ f : X \rightarrow X$ and $f \circ g : Y \rightarrow Y$ are homotopic to the corresponding identity maps $\text{id}_X : X \rightarrow X$ and $\text{id}_Y : Y \rightarrow Y$.

We remind the classical result that $\chi(X)$ is a homotopy invariant of X .

Proposition 2.4. *If simplicial complexes X, Y have the same homotopy type, then $\chi(X) = \chi(Y)$, so the Euler characteristic is a homotopy invariant.*

Let X be a simplicial complex with a piecewise linear function $f : X \rightarrow \mathbb{R}$. We assume that f is given by its real values at all vertices of X . We consider only generic functions f that have distinct values at all vertices. Otherwise we symbolically perturb the values of f by replacing the test $f(v_i) < f(v_j)$ with the test $i < j$ [11, p. 253]. Then we linearly extend f to any simplex (hence to the whole simplicial complex X) spanned by vertices v_0, \dots, v_d as follows: $f(\sum_{i=0}^d t_i v_i) = \sum_{i=0}^d t_i f(v_i)$, where all $t_i \geq 0$ and $\sum_{i=0}^d t_i = 1$.

Definition 2.5. Let a simplicial complex X have a piecewise-linear function $f : X \rightarrow \mathbb{R}$ given by distinct values at all vertices of X . The *Reeb graph* $\text{Reeb}(f, X)$ is the quotient of X , where we collapse each connected component of $f^{-1}(t) \subset X$ to a single point. Different connected components are collapsed to different points. Any point $p \in \text{Reeb}(f, X)$ represents a connected component of $f^{-1}(t)$ and has the associated value $f(f^{-1}(t)) = t$. Then we have the associated function $f_{\text{Reeb}(X)} : \text{Reeb}(f, X) \rightarrow \mathbb{R}$.

A value $f(s) \in \mathbb{R}$ is called *regular* if all level sets $f^{-1}(t)$ have the same number of connected components over a small interval $s - \delta < t < s + \delta$. Otherwise the value $f(s)$ is *critical*. A point in $\text{Reeb}(f, X)$ is a *Reeb node* if the corresponding level set component passes through a critical value of f .

An arc of $\text{Reeb}(f, X)$ is a pair (u, v) of nodes ordered by their values $f(u) < f(v)$. If $\text{Reeb}(f, X)$ has k multiple arcs between the same nodes u, v ,

then the pair (u, v) is repeated k times in the full set of arcs. We call vertices and edges of $\text{Reeb}(f, X)$ *nodes* and *arcs*, respectively, to avoid a confusion with vertices and edges of a simplicial complex. Later we shall embed all arcs of $\text{Reeb}(f, X)$ as curved (often circular) arcs in a multi-page book.

Definition 2.5 implies that $\text{Reeb}_f(X)$ has only Reeb nodes of degree 1 or at least 3, no nodes of degree 2. However, if we monitor not only connected components of a level set $f^{-1}(t)$, but also the Euler characteristic, then we need to add (topologically trivial) nodes of degree 2 to $\text{Reeb}_f(X)$ when the Euler characteristic or (more generally) any Betti number of $f^{-1}(t)$ changes.

Definition 2.6. A point $p \in \text{Reeb}_f(X)$ is called an *Euler node* of degree 2, if the corresponding level set component $C(X, p)$ changes its Euler characteristic when passing through the value $f_{\text{Reeb}(X)}(p)$. Then every arc in the new *Euler-Reeb graph* $\text{EReeb}(f, X)$ has the associated Euler characteristic.

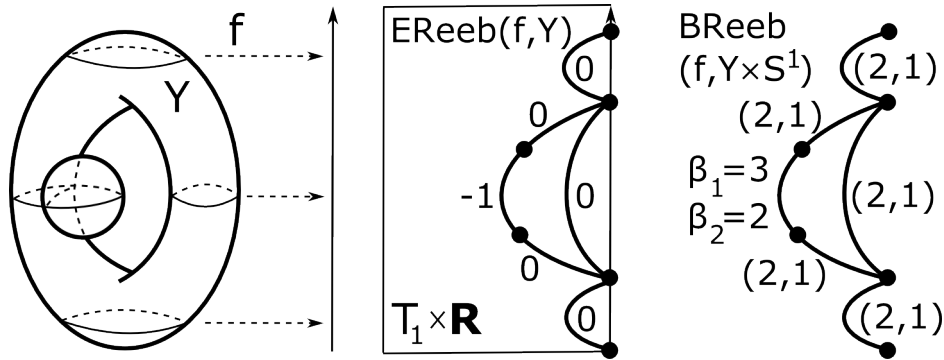


Figure 1: A torus Y with a bubble, $\text{EReeb}(f, Y)$ and $\text{BReeb}(f, Y \times S^1)$.

The middle picture in Fig. 1 shows $\text{EReeb}(f, Y)$ for a torus Y with a bubble. The bubble generates the arc with the associated Euler characteristic $\chi(\theta) = -1$, where the theta-graph θ is a component in a level set $f^{-1}(t)$.

Definition 2.7. A point $p \in \text{Reeb}_f(X)$ is called a *Betti node* of degree 2, if the corresponding level set component $C(X, p)$ changes one its Betti numbers β_i , $i \geq 1$, when passing through the value $f_{\text{Reeb}(X)}(p)$. Each β_i is the rank of the i -th homology group of $C(X, p)$. Every arc in the new Betti-Reeb graph $\text{BReeb}(f, X)$ has the associated array of Betti numbers $(\beta_1, \dots, \beta_d)$.

The right-hand side picture in Fig. 1 shows $\text{BReeb}(f, Y \times S^1)$. The bubble in Y gives a component $\theta \times S^1$ (two tori attached along an annulus) whose Betti numbers are $\beta_1 = 3$ (two meridians, one common longitude) and $\beta_2 = 2$.

3 Multi-page book embeddings of graphs

Definition 3.1. An *undirected cycle* in a graph is a sequence of distinct nodes p_1, \dots, p_n such that any nodes p_i, p_{i+1} are connected by an arc, where $i = 2, \dots, n$ and $p_{n+1} = p_1$. A *directed acyclic graph* (DAG) is a graph with oriented arcs, but without *directed cycles* of consistently oriented arcs.

Each undirected cycle in G can be considered as a formal sum of its arcs (with coefficients in $\mathbb{Z}/2\mathbb{Z}$ if we forget about orientations). Several cycles of G are called *independent* if their formal sums are linearly independent in the vector space spanned by all arcs (say with coefficients in $\mathbb{Z}/2\mathbb{Z}$ for simplicity).

Embedding Theorem 1.2 uses the maximum number $m(G)$ of independent undirected cycles that share a common arc. If cycles of G meet only at nodes or if all cycles are disjoint, they don't share any arcs, so $m(G) = 1$. The graph θ_k consisting of k arcs e_1, \dots, e_k connecting 2 nodes has $k - 1$ independent cycles $e_1 \cup e_i, i = 2, \dots, k$, which share e_1 , hence $m(\theta_k) = k - 1$.

Lemma 3.2. Any directed acyclic graph has an ordering of nodes consistent with orientations of arcs, namely any arc is oriented from a smaller node to a larger node. Any Reeb graph $\text{Reeb}(f, X)$ is a directed acyclic graph.

Proof. We can uniquely choose the partial ordering: a node u is smaller than v if there is a path of consistently oriented arcs from u to v . Then we may arbitrarily decide which of any two non-comparable nodes is larger. In $\text{Reeb}(f, X)$ we orient an arc from u to v if $f_{\text{Reeb}(X)}(u) < f_{\text{Reeb}(X)}(v)$. \square

Definition 3.3. A function $f : X \rightarrow Y$ between simplicial complexes is called an *embedding* if f is a homeomorphism on image, namely the restriction $f : X \rightarrow f(X)$ is bijective and bi-continuous (in both directions).

Definition 3.4. For any integer $k \geq 3$, if we attach k edges (*rays*) to one vertex c (*center*), we get the *star graph* T_k . The product $T_k \times \mathbb{R}$ is called the *k-page book*. The vertical axis $c \times \mathbb{R}$ is called the *binding axis* (or the *spine*). For each edge $e \subset T_k$, the product $e \times \mathbb{R}$ is a *page* of $T_k \times \mathbb{R}$, see Fig. 2.

Definition 3.5. An embedding $G \subset T_k \times \mathbb{R}$ of a directed acyclic graph G into the book is called a *book embedding* if the following properties hold:

- the *nodes-in-axis* property: all nodes of G lie in the binding axis $c \times \mathbb{R}$;
- the *directional* property: any arc has the same direction as the axis $c \times \mathbb{R}$;
- the *arc-in-one-page* property: any arc is within a single page of $T_k \times \mathbb{R}$.

The right hand side picture of Fig. 2 shows a book embedding $K_4 \subset T_2 \times \mathbb{R}$. The properties from Definition 3.5 allow us to encode a book embedding by listing all edges (say, in the increasing order of their lowest node) and by specifying the index of the page containing each arc. Fig. 1 shows more optimal embeddings of $\text{EReeb}(f, Y)$, $\text{BReeb}(f, Y \times S^1)$, not book embeddings.

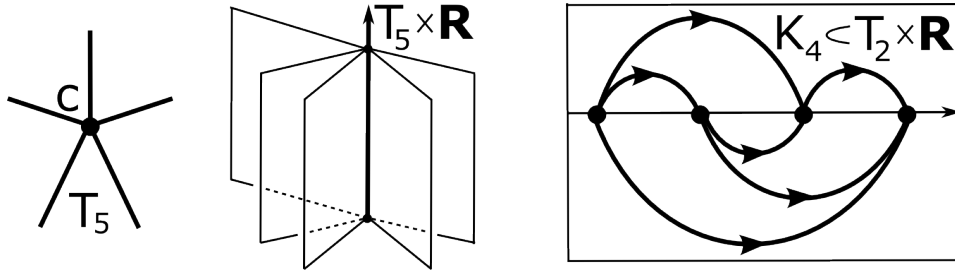


Figure 2: The star T_5 , the book $T_5 \times \mathbb{R}$, an embedding $K_4 \subset T_2 \times \mathbb{R}$.

Definition 3.6. Number nodes of a Reeb graph G by $1, 2, \dots, n$. Then a book embedding $G \subset T_k \times \mathbb{R}$ has the *code* of pairs $(ij)_l$, where $i < j$ are nodes connected by an arc in the l -th page. We may also include labels, e.g. $(ij; \chi)_l$ means that the arc (ij) has the associated Euler characteristic χ . Pairs of nodes can be lexicographically ordered if we need a unique code.

4 Embeddings of directed acyclic graphs

We were initially motivated by embeddings of Reeb graphs. However, all results in this section work for more general directed acyclic graphs.

Definition 4.1. A directed acyclic graph T is called a *tree* if T has no cycles even after forgetting all orientations. A *root* of T is a node of degree 1.

Lemma 4.2. *Any directed acyclic tree T with a root p can be embedded into the upper half-disk $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq \pm \varepsilon x, y \geq 0\}$ for any $\varepsilon > 0$ in such a way that the root p goes to the origin $(0, 0)$, all other nodes lie in the diameter on the x -axis and all arcs have the direction of the x -axis.*

Proof. We prove by induction on the number n of nodes. The base $n = 1$ is trivial. In the inductive step we may assume that the only arc e at the root p of T is oriented from p to another node q . Then we shall embed T into the positive half-disk with the center $(0.5\varepsilon, 0)$ and radius $\varepsilon > 0$. If the arc e is oriented from q to p , the symmetric construction gives a required embedding into the negative half-disk with the center $(-0.5\varepsilon, 0)$ and radius ε .

We embed the arc e as a curved arc such that the root p goes to the origin $(0, 0)$ and the node q goes to the center $(0.5\varepsilon, 0)$ of the upper half-disk. If we remove the arc e , the given tree T splits into smaller subtrees with the root q . By the inductive assumption each subtree can be embedded into a small half-disk to the left or to the right of the center $q = (0.5\varepsilon, 0)$. Fig. 3 shows yellow half-disks with similarly embedded tripods. Let the node q have ordered neighbors $q_1 < \dots < q_i < q$. We embed the subtrees of q_i one by

one starting from the lowest neighbor q_1 . We embed every next subtree into an even smaller half-disk between q and the previous subtree. We similarly embed subtrees with higher neighbors of q starting from the highest. \square

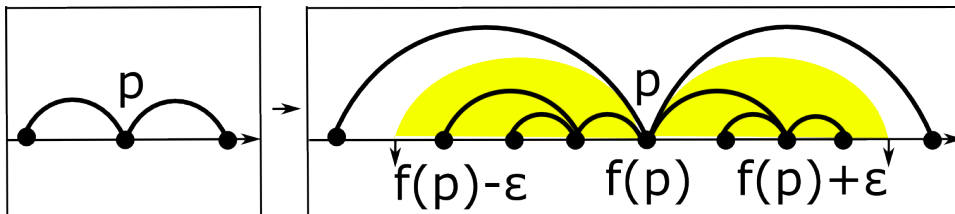


Figure 3: How to extend a book embedding by adding trees at a node p .

Any undirected graph G with n nodes has a book embedding into $\lceil \frac{n}{2} \rceil$ pages [3]. In Lemma 4.3 we embed a directed acyclic graph with an extra property that allows us to add later nodes of degree 2 and re-embed any arc in the same page by putting all new nodes into the binding axis $c \times \mathbb{R}$.

Lemma 4.3. *Any directed acyclic graph G with $n \geq 4$ ordered nodes has a canonical book embedding $G \subset T_{n-2} \times \mathbb{R}$ such that this extra property holds:*

- *the access-to-axis property: the projection to $c \times \mathbb{R}$ of any arc contains a small interval not covered by the projections of other arcs in the same page.*

Proof. We construct a required embedding $G \subset T_{n-2} \times \mathbb{R}$ by induction on the number n of nodes $p_1 < \dots < p_n$, which can be ordered by Lemma 3.2. In the base case $n = 4$ we embed $K_4 \subset T_2 \times \mathbb{R}$ as shown in Fig. 2. In the inductive step from n to $n + 1$, we take the $(n + 1)$ -st highest node p_{n+1} and embed the subgraph of G on n lower nodes by the inductive assumption.

Then we embed all arcs at p_{n+1} into the extra $(n - 1)$ -st page of the larger book $T_{n-1} \times \mathbb{R}$ as curved arcs connecting p_{n+1} to lower nodes, so we avoid any self-intersections. The embedding depends only on the order of nodes in the binding axis. If we deform the nodes keeping their order, then we can continuously deform the embedded graph G within the book $T_{n-2} \times \mathbb{R}$. \square

Definition 4.4. The *core* subgraph $\text{core}(G)$ is obtained from a directed acyclic graph G in several stages where each stage consists of these steps:

- Step (1): forget about any degree 2 node and merge its two arcs into one;
- Step (2): replace all multiple arcs between the same nodes by a single arc;
- Step (3): remove any subtree attached to a node p and keep the node p .

At each simplification stage we apply step (1), then (2), finally (3) before moving to the next stage when we again apply (1), (2), (3) and so on.

We have actually introduced the ascending filtration of the subgraphs $\text{core}(G) \subset \text{core}_1(G) \subset \cdots \subset \text{core}_k(G) = G$ when G was simplified to $\text{core}(G)$ through k stages. Namely, $\text{core}_i(G)$ is obtained from $\text{core}_{i+1}(G)$ in one stage by following Step (1), then Step (2), finally Step (3), $i = 1, \dots, k - 1$.

In fact $\text{core}(G)$ has only nodes of degree at least 3 and can be considered as a subgraph of G . So the smallest core (except a single point) is K_4 .

For the Euler-Reeb graph of the torus Y with a bubble in Fig. 1 at the first stage we (1) forget about two nodes of degree 2, (2) replace a new double arc by a single arc, (3) simplify the resulting 3-arc path to a single point.

In practice, the Reeb graph $\text{Reeb}(f, X)$ may have few cycles, because any noise in original data usually leads to short edges or trivial nodes of degree 2 that are not included in the core of the graph $\text{Reeb}(f, X)$. So the core of a Reeb graph is usually small, which leads to embeddings into small books.

Definition 4.5. For any arc e in the subgraph $\text{core}(G) \subset G$ of a directed acyclic graph G , we introduce the *edge-subgraph* $G(e) \subset G$ that was collapsed to (replaced by) the arc e in the simplification process from Definition 4.4. To make $G(e)$ unique, we order all nodes of G by Lemma 3.2 and order all arcs with the same endpoints. Let T be a tree that was attached to a node p and was removed in Step (3). We include T into $G(e)$ for the first arc e from the root p of T to the next node $q > p$ in the fixed order of nodes.

To clarify Definition 4.5, we reverse all steps of the last simplification stage from $\text{core}_1(G)$ to $\text{core}(G)$. If in Step (3) we removed a subtree T attached to a node p , then we include T into the edge-subgraph $G(e)$ for a unique edge at p . If in Step (2) we replaced k multiple arcs with the same nodes p, q by the single arc e , then $G(e)$ contains all these k arcs, so e is blown up k times. If in Step (1) we forgot about a degree 2 node in the arc e , then we simply add this node to $G(e)$. After reversing the last simplification stage, we got a larger graph $\text{core}_1(G) \supset \text{core}(G)$. Now can start building $G(e)$ for any new arc $e \subset \text{core}_1(G)$. After reversing all stages of Definition 4.5 we have constructed a unique edge-subgraph $G(e) \subset G$ for any arc $e \subset \text{core}(G)$.

A book embedding restricted to a single page is a union of disjoint curved arcs with nodes in the binding axis $c \times \mathbb{R}$. These arcs split the page into intermediate areas, see the yellow area $A(e)$ below the arc e in Fig. 4.

Lemma 4.6. *Let a directed acyclic graph G have at most m independent cycles with a common arc. Let $\text{core}(G) \subset T_{n-2} \times \mathbb{R}$ be a book embedding from Lemma 4.3. By the access-to-axis property, for any arc $e \subset \text{core}(G)$, let $I(e) \subset c \times \mathbb{R}$ be the union of intervals not covered by the projections of other arcs in the same page as e . Let $A(e)$ be the open area bounded by e , $I(e)$ and all arcs of $\text{core}(G)$ connecting the endpoints of e and $I(e)$. Let the holed*

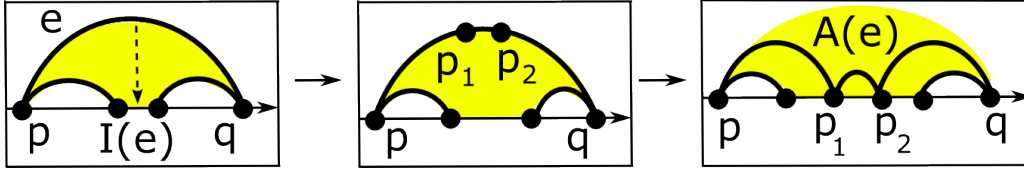


Figure 4: How to extend a book embedding after adding nodes of degree 2.

book $B(e)$ consist of $m + 1$ copies of $A(e)$ attached along $I(e)$, see the first picture in Fig. 5. Then there is a book embedding $G(e) \subset B(e) \subset T_{m+1} \times \mathbb{R}$.

Proof. Similarly to the discussion after Definition 4.5, we shall embed $G(e)$ by reversing all simplification stages from Definition 4.4. Formally we shall argue by induction on the number of stages. If $\text{core}(G) = G$, then $G(e) = e$ is already embedded. In the inductive step for the last simplification stage we use embeddings of smaller edge-subgraphs $G(e_i)$ built in fewer stages.

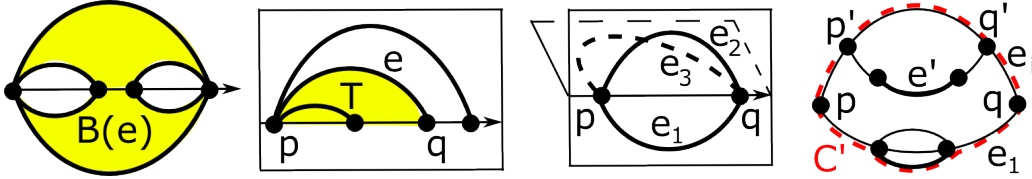


Figure 5: A holed 2-page book $B(e)$, cases of Steps (3) and (2) in Lemma 4.6.

Case of Step (3). Let the arc $e \subset \text{core}(G)$ have endpoints $p < q$ and $G(e)$ have a tree T attached at p . By Definition 4.5 the node $q > p$ is the next adjacent node to p . Hence e is the lowest arc (closest to the binding axis $c \times \mathbb{R}$) among all arcs at p in the same page. So the open area $A(e)$ looks like a half-disk bounded by e with the diameter in the binding axis, see the second picture in Fig. 5. We embed T into this half-disk by Lemma 4.2.

Case of Step (2). Let $G(e)$ contain k multiple arcs $e = e_1, \dots, e_k$ that were replaced by the arc e between nodes p, q . These k multiple arcs generate $k - 1 \leq m$ independent cycles sharing $e = e_1$. In the simplest case $G(e) = \cup_{i=1}^k e_i$ and we embed each of k arcs e_i one per page (as rotational copies of e) into the book $B(e)$ with $m + 1 \geq k$ pages, see the third picture in Fig. 5.

In general, $G(e)$ is the union of all subgraphs $G(e_i)$, $i = 1, \dots, k$, built on the arcs e_1, \dots, e_k by reversing earlier simplification stages. If each arc e_i is shared by at most m_i independent cycles in $G(e_i)$, then we need m_i extra pages to embed $G(e_i)$ by the inductive assumption. So in total we need $m_1 + \dots + m_k$ extra pages for embedding all subgraphs $G(e_i)$. However, all these $m_1 + \dots + m_k$ independent cycles from $G(e) = \cup_{i=1}^k G(e_i)$ give the same number of independent cycles that share the arc $e = e_1$ in $G(e) \subset G$.

Indeed, let each of these m_i cycles contain an arc e' , see the last picture in Fig. 5. If p', q' are extreme nodes (according to their order) in such a cycle C , we can replace C by the larger cycle C' going through $e = e_1$ (from p to q), then along e_i (from q to p) following the original cycle C through the nodes q' and p' , but avoiding e' . Then all new m_i cycles are independent in $G(e)$ and share the same subarc in e_1 as all m_1 cycles from $G(e_1)$, $i = 2, \dots, k$. So we have enough $m_1 + \dots + m_k \leq m$ pages to embed $G(e) = \cup_{i=1}^k G(e_i)$.

Case of Step (1). Let $G(e)$ have nodes p_1, \dots, p_j of degree 2 on the arc e between nodes p, q . Then we re-embed e in the same page by putting all new nodes into $I(e) \subset c \times \mathbb{R}$. See Fig. 4 for two nodes p_1, p_2 . To make this embedding canonical, we may put all p_1, \dots, p_j into the first interval from (possibly disconnected) $I(e)$ between p and the next node in $c \times \mathbb{R}$. \square

General version of Embedding Theorem 1.2. Let a directed acyclic graph G have at most m independent cycles that share a common arc. If $\text{core}(G)$ has n nodes, then G has a book embedding in $T_{(m+1)\max\{1, n-2\}} \times \mathbb{R}$.

Proof. If $\text{core}(G)$ is not a single node, then $n \geq 4$ and we take a book embedding $\text{core}(G) \subset T_{n-2} \times \mathbb{R}$ from Lemma 4.3. If $\text{core}(G)$ is a single node, then we set $n = 3$, so $\max\{1, n-2\} = 1$ for simplicity. Then G is the union of the subgraphs $G(e)$ from Definition 4.5 over all arcs $e \subset \text{core}(G)$. The open areas $A(e)$ from Lemma 4.6 are disjoint over all arcs $e \subset G$ in the same page of $T_{n-2} \times \mathbb{R}$. Hence the union of the holed books $B(e)$ over all $e \subset \text{core}(G)$ from one page of $T_{n-2} \times \mathbb{R}$ are disjoint in the blown-up $(m+1)$ -page book. For all arcs $e \subset \text{core}(G)$ in one page, we can jointly embed $G(e)$ into the disjoint union of the holed books $B(e)$ by Lemma 4.6, hence into the same $m+1$ pages. So each page of the book $T_{n-2} \times \mathbb{R} \supset \text{core}(G)$ generates $m+1$ more pages and the whole graph G is embedded into $T_{(m+1)(n-2)} \times \mathbb{R}$. \square

If we would like to canonically embed all Reeb graphs into a fixed book, we may choose a star graph with infinitely many edges $T_\infty \subset \mathbb{R}^2$ consisting of the radii of the unit circle to the points $\exp\left(\left(1 - \frac{1}{k}\right)\pi\sqrt{-1}\right)$, $k \geq 1$.

5 Discussion and further open problems

Trivial nodes of degree 2 in Reeb graphs $\text{EReeb}(f, X)$ and $\text{BReeb}(f, X)$ do not create any problem for encoding book embeddings. Indeed, we may encode a path between any nodes of degree not equal to 2 by a linear sequence of arcs with all necessary labels. All extra nodes of degree 2 in Reeb graphs have no attached subtrees. Then we may leave all multiple arcs between the same nodes in one page. For instance, the Euler-Reeb graph $\text{EReeb}(f, Y)$ for a torus Y with a bubble in Fig. 1 can be encoded by the array of arcs

$(12; 0)_1, (23; 0)_1, (23; 0, -1, 0)_1, (34; 0)_1$, where 1, 2, 3, 4 are ordered nodes and the subscript 1 highlights only 1 page. One of the arcs between nodes 2 and 3 splits into 3 subarcs with the associated Euler characteristics 0, $-1, 0$.

If we drop the requirement for a book embedding in Definition 3.5 that all nodes of degree 2 in Reeb graphs lie in the binding axis, then we strengthen Theorem 1.2 and get optimal embeddings into a 1-page book as in Fig. 1.

Corollary 5.1. *Let a Reeb graph G have at most m independent cycles that share a common arc after the first simplification stage in Definition 4.4. If $\text{core}(G)$ has n nodes, then G has an embedding into the smaller book $T_{(m+1)\max\{1, n-2\}} \times \mathbb{R}$, which can be linearly encoded as in Definition 3.6.*

The number m of independent cycles in Corollary 5.1 is smaller for a Reeb graph than in general Theorem 1.2, because G may lose multiple arcs after Steps (1) and (2) of the first simplification stage of Definition 4.4. So the number m cycles is now counted for a graph simpler than G . For instance, the Euler-Reeb graph $G = \text{EReeb}(f, Y)$ for a torus Y with a bubble in Fig. 1 initially has $m = 1$ (one cycle). However G is simplified to a single node in one stage, so $m = 0$ in Corollary 5.1 gives an embedding $G \subset T_1 \times \mathbb{R}$.

The problem to find the minimum page number $\text{PN}(G)$ such that a directed acyclic graph G has a book embedding into $T_{\text{PN}(G)} \times \mathbb{R}$ is NP-complete [7]. Our key result is an algorithmic method to embed any directed acyclic graph G into a small book, which gives a practical upper bound for $\text{PN}(G)$. The following problem is the next step for automatic processing Reeb graphs.

Problem 5.2. *Design a fast algorithm to compute a book embedding for a given directed acyclic graph, hence compute a linear code for any Reeb graph.*

Different shapes may have identical Reeb graphs. The ultimate aim is to design a shape descriptor that uniquely represents geometric objects up to a natural equivalence. We start from the homotopy equivalence on complexes.

Problem 5.3. *In addition to Betti numbers on edges of a Reeb graph, find extra invariants that would enable us to reconstruct a simplicial complex X up to (say) a homotopy equivalence from the given graph $\text{Reeb}(f, X)$.*

It is interesting to translate transformations of time-varying Reeb graphs [2] into codes of book embeddings. The $O(m \log m)$ algorithm [4, 5, 10] for computing the classical Reeb graph $\text{Reeb}(f, X)$ can probably be adapted for computing the Euler-Reeb graph. For 3-dimensional scalar fields [11] when level sets are 2-dimensional isosurfaces, all Betti numbers from $\text{BReeb}(f, X)$ were found without increasing the overall computational complexity.

Problem 5.4. *For a given simplicial complex with a real-valued function, compute the Reeb graph with all extra invariants needed for Problem 5.3.*

We are open to collaboration on the above (and any related) problems. We thank in advance any reviewers for critical comments and suggestions.

References

- [1] S. Biasotti, D. Giorgi, M. Spagnuolo, B. Falcidieno. Reeb graphs for shape analysis and applications. *Theor. Comp. Sci.* 392 (2008) 5-22.
- [2] H. Edelsbrunner, J. Harer, A. Mascarenhas, V. Pascucci, J. Snoeyink. Time-varying Reeb graphs for continuous space-time data. *Computational Geometry* 41 (2008), 149-166.
- [3] Y.-C Chen, H.-L. Fu, I.-F. Sun. Embedding graphs in books: a layout problem with applications to VLSI design. *SIAM J. Algebraic Discrete Methods* 8 (1987) 33-58.
- [4] H. Doraiswamy, V. Natarajan. Output-Sensitive Construction of Reeb Graphs. *IEEE Transactions on Visualization and Computer Graphics* 18 (2012) 146-159.
- [5] W. Harvey, Y. Wang, R. Wenger. A randomized $O(m \log m)$ time algorithm for computing Reeb graphs of arbitrary simplicial complexes. In *Proceedings of SoCG 2010*, 267-276.
- [6] L. Heath, S. Pemmaraju, A. Trenk. Stack and Queue Layouts of Directed Acyclic Graphs. Part I. *SIAM J. Comp.*, 28, 1510-1539, 1999.
- [7] L. Heath, S. Pemmaraju. Stack and Queue Layouts of Directed Acyclic Graphs. Part II. *SIAM J. Comp.*, 28, 1588-1626, 1999.
- [8] V. Kurlin. Basic embeddings into a product of graphs. *Topology and Its Applications* 102 (2000) 113-137.
- [9] V. Kurlin. Three-page encoding and complexity theory for spatial graphs. *J Knot Theory Ramifications* 16 (2007) 59-102.
- [10] S. Parsa. A deterministic $O(m \log m)$ time algorithm for the Reeb graph. In *Proceedings of SoCG 2012*, 269-276.
- [11] V. Pascucci, K. Cole-McLaughlin. Parallel computation of the topology of the level sets. *Algorithmica* 38 (2004) 249-268.
- [12] M. Yannakakis. Embedding planar graphs in four pages. *J. Comput. System Sci.* 38 (1986) 36-67.