1

ELASTIC GRAPHS

DYLAN P. THURSTON

Department of Mathematics, Indiana University, Bloomington, IN 47405, USA;
email: dpthurst@indiana.edu

Abstract

An *elastic graph* is a graph with an elasticity associated to each edge. It may be viewed as a network made out of ideal rubber bands. If the rubber bands are stretched on a target space there is an *elastic energy*. We characterize when a homotopy class of maps from one elastic graph to another is *loosening*, that is, decreases this elastic energy for all possible targets. This fits into a more general framework of energies for maps between graphs.

Contents

1. Introduction

1.1. Broader context.   We start with a general question.

QUESTION 1.1. When is one network of elastic bands looser than another?

This question could have been asked long ago, as soon as Hooke's Law was discovered. Informally, one network is looser than another if, however they are stretched, the first network will always have lower potential energy and force on the external vertices than the second. (We will make this question precise shortly, but note for now that our elastic bands have a resting length of 0, unlike real elastics or springs.)

There is a closely related question for resistors:

QUESTION 1.2. When does one network of resistors dissipate more energy than another?

Elastic bands and resistors both have quadratic responses to their inputs: the potential energy stored in an elastic band is a quadratic function of how far it is stretched, while the energy dissipated in a resistor is a quadratic function of the voltage difference between the ends. (See Appendix B for more on this relationship.)

The electrical version of the question is much easier. Consider a network of resistors with k external *nodes* where the voltage can be controlled. Such a network has a symmetric $k \times k$ *response matrix*, giving the vector of currents flowing out of each node as a linear function of the vector of voltages. This same matrix also gives a positive semidefinite quadratic form giving the total energy dissipated, so Question 1.2 amounts to asking whether one quadratic form dominates another.

This answer does not translate to Question 1.1. The key difference between resistors and elastics is that electricity has a direction of flow through a resistor, while elasticity just has an unsigned tension. Thus in a resistor network the balancing condition at a vertex (Kirchhoff's first law) is a linear equation, while in an elastic network, stretched on a target space that is also a graph, there are some combinatorics that play a role as well. The difference between Questions 1.2 and 1.1 can be compared to the difference between homology and homotopy, or between holomorphic differentials and quadratic differentials.

In this paper we answer one interpretation of Question 1.1. The answers are different than for Question 1.2; for instance, the classical $Y-\Delta$ transform [**Ken99**], giving an equivalence between different resistor networks, becomes an inequality;

cf. (B.3). The answer is in terms of minimizing a new 'embedding energy' for maps between graphs; cf. Equation (1.15).

There has been a great deal of work related to the electrical networks version of the problem (or, equivalently, harmonic maps from a graph to $\mathbb{R}$). Most prominently, since a harmonic $\mathbb{R}$-valued function on a simply connected planar domain is the real part of an analytic $\mathbb{C}$-valued function, it can be used to get discrete analogues of analytic functions. There is a 70-year history of work on the discrete harmonic or analytic functions [**BSST40**, **Isa41**, **Fer44**, **Duf68**, **PP93**, **Mer01**, **Smi01**, **Lov04**, **CS11**, *inter alia*]. See Smirnov's ICM lecture [**Smi10**] for some of the history. These works rely crucially on the target being a vector space. It is, for instance, not clear how to apply these discretizations to compute the minimal dilatation map between Riemann surfaces.

There is another line of work toward discrete approximations to analytic functions, based on circle packings on the Riemann sphere, as proposed by Thurston [**Thu86**] and developed by others [**RS87**, **HS96**, **Ste05**]. Here the target of the map is the sphere, not a vector space, but it does use a complex projective structure on the target to define the meaning of a 'circle' [**KMT03**]. The method also involves transcendental equations.

In this paper, we are interested in more general targets. Concretely, we work with maps to graphs rather than to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{CP}^1$. The equations that result have combinatorial data and, once the combinatorics are fixed, are algebraic. For instance, for harmonic maps from one graph to another, the equations are linear once the combinatorial data is fixed. Although we work mainly with targets that are graphs, in fact the main results apply to energies of harmonic maps with *any* possible target.

See Section 1.4 for more on prior work.

Although in principle the results are quite general, the main applications to date are in the setting of surfaces. We briefly summarize the connection to surfaces here, although in the bulk of the paper we will not refer to it. A *ribbon graph* is a graph together with a cyclic ordering of the edges incident to each vertex; this is enough to give a topological thickening of the graph to a surface. In the setting of surfaces, it is traditional to look not at harmonic energy but at a dual notion, *extremal length*. This is based on homotopy classes of loops mapped into the surface, rather than on maps from the surface to some other space. (Definition 1.19 has the corresponding notion for graphs.) In the graph setting, the control we get over when one network is looser than another also gives bounds on the maximal ratio of extremal lengths, maximized over all homotopy classes of loops. This turns out to be quite useful, because for closed surfaces, the maximum ratio of extremal lengths over all homotopy classes gives the minimal dilation of a quasiconformal map between the surfaces [**Ker80**, Theorem 4], while for surfaces

with boundary, this ratio controls whether one surface conformally embeds in another [**KPT15**].

In turn, this control over when one surface conformally embeds in another (plus a relation between extremal length on a graph and extremal length on its thickening) lets us prove a new characterization of when a topological branched self-cover of the sphere is equivalent to a rational map [**Thu16b**]. This gives a converse to an earlier characterization by Thurston [**DH93**].

You can also use these techniques and the relation between graphs and surfaces to approximate the Teichmüller distance between two Riemann surfaces [**Pal15**]. In principle, that work does not use the results of this paper, since we need a more flexible notion of 'looser' than the one considered here. But the motivation and techniques are similar.

1.2. Definitions and results.

We now turn to making Question 1.1 precise.

DEFINITION 1.3. In this paper, a *graph* is a topologist's graph, a finite 1-dimensional CW complex, that is, with multiple edges and self-loops allowed. *Maps* between graphs are continuous maps between the underlying topological spaces. In particular, they need not send vertices to vertices. For convenience, we will work with maps that are piecewise-linear (PL) with respect to a fixed linear structure on each edge.

DEFINITION 1.4. A *marked graph* is a pair (Γ, M), where Γ is a graph and $M \subset \text{Vert}(\Gamma)$ is a finite subset of marked points (possibly empty). A map $f : (\Gamma_1, M_1) \to (\Gamma_2, M_2)$ between marked graphs is required to send marked points to marked points (that is, $f(M_1) \subset M_2$). Homotopy is considered within the space of such maps. In particular, within a homotopy class the restriction of f to a map from M_1 to M_2 is fixed.

DEFINITION 1.5. A *length graph* $K = (\Gamma, \ell)$ is a graph in which each edge e has a positive length $\ell(e)$. This gives a metric on Γ respecting the linear structure on the edges. If $f : K_1 \to K_2$ is a PL map between length graphs, then $|f'| : K_1 \to \mathbb{R}_{\geqslant 0}$ is the absolute value of the derivative with respect to the metrics. It is locally constant with jump discontinuities.

DEFINITION 1.6. An *elastic graph* $G = (\Gamma, \alpha)$ is a graph Γ in which each edge e has a positive elastic constant $\alpha(e)$. We can take derivatives by interpreting $\alpha(e)$ as the length of e as above. For G an elastic graph, K a length graph, and $f : G \to K$ a PL map, the *Dirichlet energy* of f is

$$\mathrm{Dir}(f) := \int_{x \in \Gamma} |f'(x)|^2 \, dx. \tag{1.7}$$

The Dirichlet energy of a (marked) homotopy class $[f]$ is defined to be

$$\mathrm{Dir}[f] := \inf_{g \in [f]} \mathrm{Dir}(g). \tag{1.8}$$

It is easy to see that a minimizer g for $\mathrm{Dir}[f]$ must be *constant-derivative*: $|g'|$ is constant on each edge of g (Definition 4.7). If g is constant-derivative, we then have

$$\mathrm{Dir}(g) = \sum_{e \in \mathrm{Edge}(\Gamma)} \frac{\ell(g(e))^2}{\alpha(e)}, \tag{1.9}$$

where $\ell(g(e))$ is the length of the image of e in a natural sense (Definition 2.2). This is the Hooke's law energy of g, where each edge is an ideal spring with resting length 0 and spring constant given by α. Physically, you could think about stretching a rubber band graph shaped like G over a system of pipes shaped like K.

Harmonic maps are maps that minimize the energy (1.7) or (1.9) within their homotopy class. Intuitively, a map is harmonic if it is constant-derivative and each vertex is at equilibrium, in the sense that the force pulling in one direction is less than or equal to the total force pulling it in other directions. Physically, the force pulling a vertex in the direction of an incident edge e is given by the derivative of Equation (1.9) with respect to $\ell(g(e))$, that is, $2\ell(g(e))/\alpha(e) = 2|f'(e)|$. We will drop the irrelevant factor of 2 and refer to $|f'(e)|$ as the *tension* in the edge (Definition 5.1).

For instance, if a vertex v of G maps to the middle of an edge of K with three edges on the left and one edge on the right as on top of Figure 1, the net force to the left must equal the net force to the right:

$$|f'(a)| + |f'(b)| + |f'(c)| = |f'(d)|.$$

On the other hand, if a vertex v of G maps to a vertex w of K as on the bottom of Figure 1, we no longer have an equality. Instead, the net force pulling v into any one of the three edges incident to w cannot be greater than the total force pulling it into the other two edges. This gives three triangle inequalities:

$$\begin{aligned}
|f'(a)| + |f'(b)| &\leqslant |f'(c)| + |f'(d)| \\
|f'(c)| &\leqslant |f'(a)| + |f'(b)| + |f'(d)| \\
|f'(d)| &\leqslant |f'(a)| + |f'(b)| + |f'(c)|.
\end{aligned} \tag{1.10}$$

(See Corollary 5.3 for the general statement.) In any homotopy class of maps to a target length graph there is a harmonic map (Theorem 5).
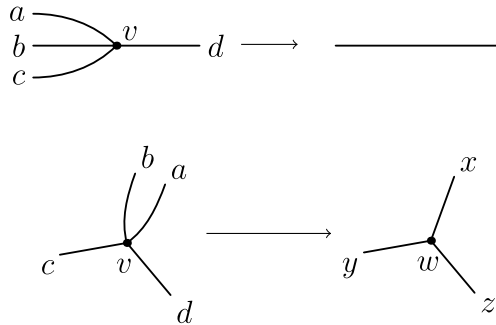
Figure 1. Local schematic pictures of the equilibrium condition for harmonic maps. Top: A vertex maps to an edge. Bottom: A vertex maps to a vertex.

EXAMPLE 1.11. Consider an elastic graph G and length graph K, both tripods with their ends marked, with elastic constants and lengths given by
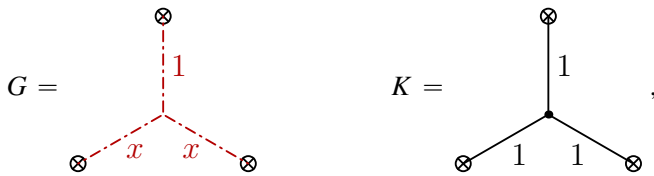
where x varies. If $x = 3$, the map f minimizing Dirichlet energy takes the vertex of G to a point $1/5$ of the way along an edge:
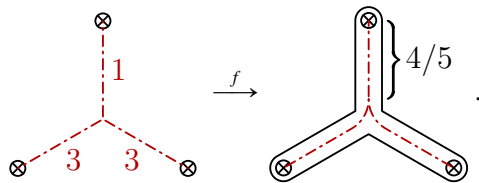
The net force pulling the vertex of G upwards is $4/5$, the same as the net force pulling it downwards. On the other hand, if $0 < x \leqslant 2$, then the harmonic representative has the central vertex of G mapping to the central vertex of K.

We can think of Dirichlet energy as a function of the target lengths (or more precisely the target lengths and the homotopy class). We next compare these functions.

DEFINITION 1.12. Given a homotopy class $[\phi]: G_1 \to G_2$ of maps between marked elastic graphs, we say that $[\phi]$ is *loosening* if, for all marked length graphs K and marked maps $f: G_2 \to K$,

$$\mathrm{Dir}[f \circ \phi] \leqslant \mathrm{Dir}[f].$$

If the inequality is always strict, we say that $[\phi]$ is *strictly loosening*. For a finer invariant, define the *Dirichlet stretch factor* to be

$$\mathrm{SF}_{\mathrm{Dir}}[\phi] := \sup_{[f]: G_2 \to K} \frac{\mathrm{Dir}[f \circ \phi]}{\mathrm{Dir}[f]}, \tag{1.13}$$

where the supremum runs over all marked length graphs K and all marked homotopy classes $[f]$, so that $[\phi]$ is loosening if and only if $\mathrm{SF}_{\mathrm{Dir}}[\phi] \leqslant 1$. It is less obvious but also true that $[\phi]$ is strictly loosening if and only if $\mathrm{SF}_{\mathrm{Dir}}[\phi] < 1$.

DEFINITION 1.14. For $\phi: G_1 \to G_2$ a PL map between marked elastic graphs, the *embedding energy* of ϕ is

$$\mathrm{Emb}(\phi) := \operatorname*{ess\,sup}_{y \in G_2} \sum_{x \in \phi^{-1}(y)} |\phi'(x)|. \tag{1.15}$$

We are taking the essential supremum over points $y \in G_2$, ignoring in particular vertices of G_2, images of vertices of G_1, and points where $\phi^{-1}(y)$ is infinite, all of which have measure zero. For a homotopy class $[\phi]$, define

$$\mathrm{Emb}[\phi] := \inf_{\psi \in [\phi]} \mathrm{Emb}(\psi). \tag{1.16}$$

EXAMPLE 1.17. If G_1 and G_2 are both tripods with marked ends, with elastic constants $(1, 3, 3)$ and $(1, 1, 1)$, then the minimizer for $\mathrm{Emb}(\phi)$ is not the map from Example 1.11, but instead sends the vertex of G_1 to a point $2/5$ of the way along an edge of G_2, with $\mathrm{Emb}(\phi) = 3/5$:
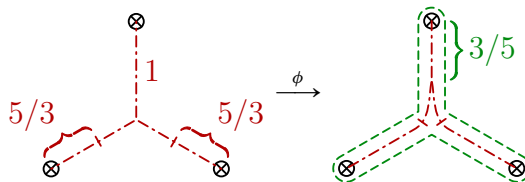
Our answer to Question 1.1 says that $\mathrm{Emb}[\phi] = \mathrm{SF}_{\mathrm{Dir}}[\phi]$, and in particular the homotopy class $[\phi]$ is loosening exactly when $\mathrm{Emb}[\phi] \leqslant 1$. Before stating the theorem, we give another way to measure the tightness of an elastic network G, using maps to G rather than maps from G.

DEFINITION 1.18. A *marked one-manifold* is a marked graph in which every marked point is a 1-valent vertex and every other vertex is 2-valent; equivalently, it is a (not necessarily connected) one-manifold C with boundary, where the set of marked points is equal to ∂C. A *marked multicurve* on a marked graph Γ is a marked one-manifold C and a marked map $c\colon C \to \Gamma$. Thus, it is a union of loops on Γ and arcs between marked points of Γ. It is a *marked curve* if C is connected. A marked multicurve is *reduced* if it is PL and has no backtracking: on each component of C, c is either constant or has a perturbation that is locally injective. In each homotopy class of multicurves there is an essentially unique reduced representative. If (C, c) is a marked multicurve on Γ, then for $y \in \Gamma$, define $n_c(y)$ to be the size of $c^{-1}(y)$. If c is reduced and e is an edge of Γ, then $n_c(y)$ is constant almost everywhere on e, so we may write $n_c(e)$.

A *marked weighted multicurve* is a marked multicurve in which each component C_i has a positive weight w_i. For weighted multicurves, $n_c(y)$ is the weighted count of $c^{-1}(y)$.

DEFINITION 1.19. The *extremal length* of a reduced multicurve c on an elastic graph $G = (\Gamma, \alpha)$ is

$$\mathrm{EL}(c) := \sum_{e \in \mathrm{Edge}(G)} \alpha(e) n_c(e)^2. \tag{1.20}$$

This is also the extremal length of the homotopy class $[c]$; see Equation (2.5) for extremal length of nonreduced multicurves.

See [Thu16a, Section 5.2] for motivation on why this is called extremal length.

DEFINITION 1.21. For $[\phi]\colon G_1 \to G_2$ a homotopy class of maps between elastic marked graphs, the *extremal length stretch factor* is

$$\mathrm{SF}_{\mathrm{EL}}[\phi] := \sup_{[c]\colon C \to G_1} \frac{\mathrm{EL}[\phi \circ c]}{\mathrm{EL}[c]} \tag{1.22}$$

where the supremum runs over all homotopy classes of marked multicurves (C, c) on G_1.

We are now ready to state the main result of this paper.

THEOREM 1. *For* $[\phi]\colon G_1 \to G_2$ *a homotopy class of maps between marked elastic graphs,*

$$\mathrm{Emb}[\phi] = \mathrm{SF}_{\mathrm{Dir}}[\phi] = \mathrm{SF}_{\mathrm{EL}}[\phi]. \tag{1.23}$$

Furthermore, there is:

- *a PL map $\psi \in [\phi]$ realizing* Emb$[\phi]$;

- *a marked length graph K and a PL map $f : G_2 \to K$ realizing* SF$_{\mathrm{Dir}}$, *in the sense that f and $f \circ \phi$ are both harmonic and*

$$\mathrm{SF}_{\mathrm{Dir}}[\phi] = \frac{\mathrm{Dir}(f \circ \psi)}{\mathrm{Dir}(f)};$$

and

- *a marked weighted multicurve (C, c) on G_1 realizing* SF$_{\mathrm{EL}}$, *in the sense that c and $\psi \circ c$ are both reduced and*

$$\mathrm{SF}_{\mathrm{EL}}[\phi] = \frac{\mathrm{EL}(\psi \circ c)}{\mathrm{EL}(c)}.$$

Note that Emb$[\phi]$ is defined as an infimum (over the homotopy class), while SF$_{\mathrm{Dir}}[\phi]$ and SF$_{\mathrm{EL}}[\phi]$ are defined as suprema (over all possible targets or multicurves). As such, Equation (1.23) helps us compute these quantities, by sandwiching the target value.

We also give an algorithm that produces, simultaneously, the map $\psi \in [\phi]$ minimizing Emb, the pair (K, f) maximizing the ratio of Dirichlet energies, and the reduced multicurve (C, c) maximizing the ratio of extremal lengths.

EXAMPLE 1.24. Example 1.17 fits into the following sequence of maps realizing SF$_{\mathrm{EL}}$ and SF$_{\mathrm{Dir}}$:

$$C \xrightarrow{c} G_1 \xrightarrow{\phi} G_2 \xrightarrow{g} K$$
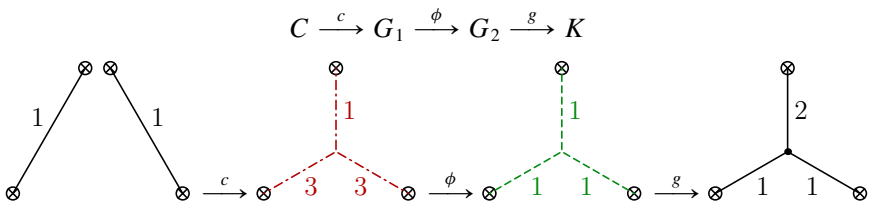
Here c is the evident map, ϕ is the map from Example 1.17, and g is the map that takes the vertex of G_2 to the vertex of K. Specifically, we have

$$\frac{\mathrm{EL}[\phi \circ c]}{\mathrm{EL}[c]} = \frac{4 + 1 + 1}{4 + 3 + 3} = \mathrm{Emb}(\phi) = 3/5 = \frac{\mathrm{Dir}[g \circ \phi]}{\mathrm{Dir}[g]}$$

$$= \frac{36/25 + 27/25 + 27/25}{4 + 1 + 1}.$$

We can generalize the target spaces in Theorem 1 considerably.

DEFINITION 1.25. Let G be a marked elastic graph and let X be a marked length space. (A *length space* is a metric space where the distance is given by the infimum of lengths of paths. *Marked* means that there is a distinguished finite set of points.) Let $f\colon G \to X$ be a Lipschitz map taking marked points to marked points. Define the Dirichlet energy of f by

$$\operatorname{Dir}(f) := \int_{x \in G} |f'(x)|^2 \, dx$$

where $|f'(x)|$ is the best local Lipschitz constant of f in a neighborhood of x. For a homotopy class of marked maps, define

$$\operatorname{Dir}[f] := \inf_{g \in [f]} \operatorname{Dir}(g).$$

In this generality, minimizers for $\operatorname{Dir}[\phi]$ need not exist. (See [**KS93**, **EF01**] for some cases where minimizers do exist.)

THEOREM 2. *For $[\phi]\colon G_1 \to G_2$ a homotopy class of maps between marked elastic graphs,*

$$\operatorname{Emb}[\phi] = \sup_{[f]\colon G_2 \to X} \frac{\operatorname{Dir}[f \circ \phi]}{\operatorname{Dir}[f]},$$

where the supremum runs over all marked length spaces X and all homotopy classes of marked maps $f\colon G_2 \to X$ with $\operatorname{Dir}[f] > 0$.

We next put Theorem 1 in the broader context of a family of energies of maps between weighted graphs, elastic graphs, and length graphs:

$$(1.26)$$
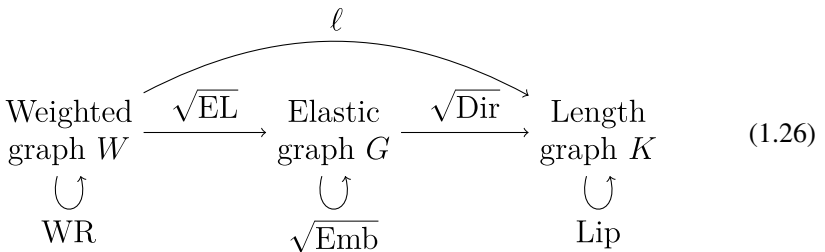
(Weighted graphs, Definition 2.3, are a mild generalization of weighted multicurves.) The label on an arrow gives the appropriate energy of a map between the given types of graphs. $\operatorname{EL}(f)$, $\operatorname{Emb}(f)$ and $\operatorname{Dir}(f)$ are as defined

above, except that we take the square root for reasons to be explained shortly, and EL has been extended from multicurves to maps from general weighted graphs in a natural way (Equation (2.5)). We make some additional definitions.

- For $f \colon W \to K$ a PL map from a weighted graph to a length graph, $\ell(f)$ is the weighted length of the image of f:

$$\ell(f) := \int_{x \in W} w(x)|f'(x)| \, dx. \tag{1.27}$$

- For $f \colon K_1 \to K_2$ a PL map between length graphs, $\mathrm{Lip}(f)$ is the best global Lipschitz constant for f:

$$\mathrm{Lip}(f) := \operatorname*{ess\,sup}_{x \in K_1} |f'(x)|. \tag{1.28}$$

- For $f \colon W_1 \to W_2$ a PL map between weighted graphs, the *weight ratio* is the maximum ratio of weights:

$$\mathrm{WR}(c) := \operatorname*{ess\,sup}_{y \in W_2} \frac{\sum_{x \in f^{-1}(y)} w(x)}{w(y)}. \tag{1.29}$$

A key fact is that these energies are *submultiplicative*, in the sense that the energy of a composition of two maps is less than or equal to the product of the energies of the two pieces. To state this uniformly, we make some further definitions.

DEFINITION 1.30. For $p \in \{1, 2, \infty\}$, a *p-conformal graph* G^p is:

- for $p = 1$, a weighted graph;

- for $p = 2$, an elastic graph; and

- for $p = \infty$, a length graph.

Let $p, q \in \{1, 2, \infty\}$ with $p \leqslant q$. If we have a p-conformal graph G_1, a q-conformal graph G_2, and a PL map $f \colon G_1^p \to G_2^q$ between them, define $E_q^p(f)$ by

$$E_1^1(f) := \mathrm{WR}(f) \qquad E_2^1(f) := \sqrt{\mathrm{EL}(f)} \qquad E_\infty^1(f) := \ell(f)$$
$$E_2^2(f) := \sqrt{\mathrm{Emb}(f)} \qquad E_\infty^2(f) := \sqrt{\mathrm{Dir}(f)}$$
$$E_\infty^\infty(f) := \mathrm{Lip}(f).$$

In each case, for a homotopy class $[f]$ of maps, define

$$E_q^p[f] := \inf_{g \in [f]} E_q^p(g).$$

Let $p, q, r \in \{1, 2, \infty\}$ with $p \leqslant q \leqslant r$. Suppose we have a sequence of maps

$$G_1^p \xrightarrow{f} G_2^q \xrightarrow{g} G_3^r$$

between marked graphs of the respective types. Then the energies are submultiplicative:

$$E_r^p(g \circ f) \leqslant E_q^p(f) E_r^q(g) \tag{1.31}$$

$$E_r^p[g \circ f] \leqslant E_q^p[f] E_r^q[g]. \tag{1.32}$$

See Proposition 2.15 for a more detailed statement, spelling out the cases.

More generally, p-conformal graphs and energies E_q^p can be defined uniformly for real numbers p, q with $1 \leqslant p \leqslant q \leqslant \infty$. See Appendix A.

Theorem 1 can be interpreted as saying that some cases of Equation (1.32) are tight, as follows.

DEFINITION 1.33. Let $p, q \in \{1, 2, \infty\}$ with $p \leqslant q$. If $f \colon G_1^p \to G_2^q$ is a PL map between marked graphs of the respective conformal type and

$$E_q^p(f) = E_q^p[f],$$

we say that f is an *energy minimizer* (in its homotopy class).

Let $p, q, r \in \{1, 2, \infty\}$ with $p \leqslant q \leqslant r$. If we have a sequence of maps

$$G_1^p \xrightarrow{f} G_2^q \xrightarrow{g} G_3^r$$

of PL maps between graphs of the respective conformal type, we say that this sequence is *tight* if $g \circ f$ is an energy minimizer and Equation (1.31) is sharp:

$$E_r^p[g \circ f] = E_r^p(g \circ f) = E_q^p(f) E_r^q(g).$$

We may similarly say that a longer sequence of maps is tight.

LEMMA 1.34. *Let* $G_1^p \xrightarrow{f} G_2^q \xrightarrow{g} G_3^r$ *be a sequence of maps. If the sequence is tight, then* f *and* g *are also energy minimizers. Conversely, if* f *and* g *are energy minimizers and* $E_r^p[g \circ f] = E_q^p[f] E_r^q[g]$, *then the sequence is tight.*

Proof. In any sequence of maps $G_1^p \xrightarrow{f} G_2^q \xrightarrow{g} G_3^r$ (tight or not), we have the following inequalities:

$$E_r^p[g \circ f] \leqslant E_r^p(g \circ f) \leqslant E_q^p(f) E_r^q(g)$$

$$E_r^p[g \circ f] \leqslant E_q^p[f] E_r^q[g] \leqslant E_q^p(f) E_r^q(g).$$

(In both lines, we use Proposition 2.15 and the definition of $E_q^p[f]$ as an infimum.) The hypothesis of the first (respectively second) claim in the statement is that the inequalities in the top (respectively bottom) line are equalities. In either case, the inequalities in the other line must be equalities as well. □

In this language, Theorem 1 says that, for every homotopy class of maps $[\phi]\colon G_1 \to G_2$ between marked elastic graphs, there is a corresponding tight sequence

$$C \xrightarrow{c} G_1 \xrightarrow{\psi} G_2 \xrightarrow{f} K$$

with $\psi \in [\phi]$, (C, c) a weighted multicurve on G_1, and K a length graph.

1.3. Structure of the paper. The paper is organized by the type of energy under consideration. In the course of the paper, we also prove several other tightness results.

- In Section 3, we prove Theorem 3, which gives a version of the max-flow/min-cut theorem in the context of maps between graphs. This gives a characterization of which maps minimize E_p^1 for any p. This is surprisingly involved, and feels like it should be standard, but we have been unable to find a reference.

- In Section 4, we prove Theorem 4, which shows that for any homotopy class $[\phi]\colon K_1 \to K_2$ of maps between marked length graphs, there is a tight sequence

$$C \xrightarrow{c} K_1 \xrightarrow{\psi} K_2$$

where $\psi \in [\phi]$. This is a theorem of White [**Bes11**]: the minimal Lipschitz stretch factor between metric graphs equals the maximal ratio by which lengths of multicurves are changed. We also extend the graphs we are looking at to allow *weak graphs*, where lengths are allowed to be zero, and prove the corresponding tightness result (Theorem 4′).

- In Section 5, we prove Theorem 5, which shows that for any homotopy class $[f]\colon G \to K$ of maps from a marked elastic graph to a marked length graph, there is a tight sequence

$$C \xrightarrow{c} G \xrightarrow{g} K$$

where $g \in [f]$ is harmonic. Again, we extend the results to the setting of weak graphs (Theorem 5′).

- Section 6 has the proof of the main results, Theorems 1 and 2, characterizing minima Emb(ϕ) in a homotopy class.

- Appendix A defines a family of energies $E_q^p(f)$ for any $1 \leqslant p \leqslant q \leqslant \infty$ and any PL map f between metric graphs. Theorem 6 gives the most general statement: for any $1 \leqslant p \leqslant q \leqslant \infty$ and homotopy class of maps $[\phi] \colon G \to H$ between marked graphs with the appropriate extra structure, we can find $\psi \in [\phi]$ and a tight sequence

$$ C \xrightarrow{\ c\ } G \xrightarrow{\ \psi\ } H \xrightarrow{\ f\ } K. $$

(The proof of Theorem 6 is only sketched.)

- Finally, Appendix B relates the elastic networks of this paper and the more well-studied subject of resistor networks and electrical equivalence.

See Table 1 for the various structures we use on graphs and on maps between them.

Throughout the paper, we develop the theory behind these concepts, more than is necessary for the proof of the main Theorem 1. For instance, the theory of taut maps in Section 3 is more general than needed for the maps that actually arise in Sections 5 and 6.

Many of the definitions and proofs are made more lengthy by the need to deal with edges of weight or length 0. (These arise when dealing with PL maps that are constant on some edges.) To get an overall view, these extra complications are unnecessary. Thus we recommend skipping the following sections on first reading:

- Assume that all reduced maps are locally injective in Definition 3.3. This makes much of Section 3 trivial.

- The notion of *weak graphs* and *weak maps* in Section 4. With that simplification, the older Theorem 4 suffices (not the more complicated Theorem 4′), and almost the entire section can be skipped.

- Section 5.2.

- Section 6.3.

In addition, for readers familiar with Riemann surfaces and especially flip-translation surfaces, the discussion about rectangle tilings at the end of Appendix B may be helpful.

Table 1. Structures on graphs and maps in this paper. See also Definition 1.30.

Concept	Summary	Letters	Reference
Discrete structures on graphs		Γ	
Marked graph	With special marked points		Definition 1.4
Multicurve	1-manifold	C, c	Definition 1.18
Weak graph	Null edges of 0 length allowed		Definition 4.1
Train track	Edges partitioned into gates	T	Definition 3.14
Metric or numeric structures on graphs			
Weighted	With multiplicity, ∞-conformal	$W, w, \mathcal{W}$	Definition 2.3
Balanced	Weights satisfying Δ inequalities		Definition 3.14
Length	Metric graph, ∞-conformal	$K, \ell, \mathcal{L}$	Definition 1.5
Elastic	Stretchable edges, 2-conformal	G, α	Definition 1.6
Strip	Weighted & length structure	$S, (w, \ell)$	Definition 6.1
p-conformal	Generalization of above	G, α	Definition A.1
Types of maps between graphs			
Edge-reduced	No backtracking on individual edges		Definition 3.1
Reduced	No backtracking		Definition 3.3
Taut	Minimizing E_p^1 for all p		Definition 3.6
Strongly reduced	Taut for some choice of weights		Definition 3.35
Constant-derivative	$|f'|$ constant on edges		Definition 4.7
Harmonic	Minimizing E_∞^2	f	Definition 5.1
(Partially) λ-filling	Minimizing E_2^2	ϕ	Definitions 6.4, 6.10
Relaxed	Only metric information	r	Definition 5.9

1.4. Prior and related work.

In addition to the work on discrete harmonic functions surveyed at the beginning of the introduction, harmonic maps between manifolds and singular spaces have been studied for a long time, and there is a great deal of work on various cases [**GS92**, **KS93**, **Wol95**, **Pic05**, *inter alia*]. In particular, Eells and Fuglede proved that in every homotopy class of maps between suitable (marked) Riemannian polyhedra there is a harmonic representative [**EF01**, Theorem 11.1], which is a large part of our Theorem 5. Their theory is much more general. We reprove these results in our more elementary setting, while also providing further connections to, for example, the theory of extremal length.

There have also been various notion of extremal length on graphs. Duffin [**Duf62**] was one of the first, defining extremal length between two nodes in a planar graph, using definitions very similar to ours. Since they only considered two nodes, this is related to maps to $\mathbb{R}$ and the case of electrical networks, as

explained in their paper. A somewhat different notion of extremal length was considered by Cannon, Floyd, and Perry [**CFP94**], where the contributions to the 'length' part of 'extremal length' come from passing through vertices rather than over edges. Thus, even though they only consider rectangles and cylinders (both related to maps to $\mathbb{R}$), their theories are not quite equivalent to electrical networks, although they also get square tilings. Their vertex-based extremal length is coarsely related to the edge-based extremal length in the current paper, with constants depending on the maximum valence.

Question 1.1 appears to be new.

The definition of embedding energy in Definition 1.14 has not appeared in this form, although it is dual to Jeremy Kahn's notion of domination of weighted arc diagrams [**Kah06**]. See [**Thu16b**, Proposition 1.17] for the precise statement.

Theorem 1 should be thought of as analogous to Teichmüller's theorem, that in any homotopy class of homeomorphism $[f]\colon S_1 \to S_2$ between closed Riemann surfaces there are 'best' metrics on S_1 and S_2 that reflect the minimal quasiconformal stretching. (The length graph K in the statement of Theorem 1 turns out to be G_2 with a different metric.)

Most of the general graph energies in Appendix A appear to be new, but the energy $E_p^1[c]$ is a power of the p-modulus of the homotopy class, which exists in much greater generality [**Fug57**].

Many of the results of this paper were announced as part of an earlier research report [**Thu16a**], which also contains many related open problems.

2. Basic notions and submultiplicativity

As mentioned already, all graphs in this paper have a linear structure on the edges, metrics are assumed to be linear with respect to this structure, weights are assumed to be piecewise-constant, and maps between graphs are assumed to be PL. All the energies extend naturally to a wider class of maps between graphs; however, the exact wider class of maps depends on the energy, and it is more convenient to stick to PL maps. The downside of working with PL maps is that the existence of minimizers of the energies is not obvious, since the space of PL maps is not locally compact in any reasonable sense. We will prove existence of energy minimizers in each case of interest.

We start by revisiting our definitions of graphs and energies.

DEFINITION 2.1. If $f\colon \Gamma_1 \to \Gamma_2$ is a PL map, the *regular points* of f are the points in Γ_1 that are in the interior of a segment on which f is linear with nonzero derivative, and the *singular points* are all other points, namely vertices of Γ_1, preimages of vertices of Γ_2, points where the derivative changes, and segments

on which f is constant. Similarly, the *singular values* in Γ_2 are the images of singular points, and the *regular values* are the rest of Γ_2. There are only finitely many singular values. If $y \in \Gamma_2$ is any point in the domain (regular or singular), a *regular neighborhood* of y is any neighborhood that has no singular values, except possibly at y.

DEFINITION 2.2. As in the introduction, a length graph K is a pair (Γ, ℓ) of a graph and an assignment of a length $\ell(e) > 0$ to edge e of Γ; we call ℓ a *metric* on Γ. In a *weak length graph*, we weaken the conditions to $\ell(e) \geqslant 0$. The space of all weak metrics on Γ is denoted by $\mathcal{L}(\Gamma)$, and the subspace of metrics is denoted by $\mathcal{L}^+(\Gamma)$. See Definition 4.1 for more on weak metrics. Given a PL map $f: \Gamma_1 \to (\Gamma_2, \ell)$ to a length graph, we can pull-back the metric on Γ_2 to a weak metric $f^*\ell$ on Γ_1, assigning to each edge $e \in \text{Edge}(\Gamma_1)$ the total length traced out by $f(e)$. We also write $\ell(f(e))$ for $f^*\ell(e)$.

DEFINITION 2.3. A *weighted graph* is a graph $W = (\Gamma, w)$ with a piecewise-constant weight function $w: \Gamma \to \mathbb{R}_{\geqslant 0}$. If w is constant on an edge e, we write $w(e)$ for the weight on e. The space of all weights on Γ that are constant on each edge is denoted by $\mathcal{W}(\Gamma)$, and $\mathcal{W}^+(\Gamma)$ is the subspace where the weights are positive.

DEFINITION 2.4. For W a weighted graph, Γ a graph, and $c: W \to \Gamma$ a PL map, the *multiplicity function* $n_c: \Gamma \to \mathbb{R}$ is defined at regular values of c by

$$n_c(y) := \sum_{x \in c^{-1}(y)} w(x).$$

This may also be written n_c^w or c_*w if we need to make the dependence on w explicit. We will never care about the value of n_c at nonregular values.

If in addition Γ has more structure we have an energy.

- If Γ is also a weighted graph with its own weights w, the *weight ratio* was defined in Equation (1.29); it is the ℓ^∞ norm of n_c.

- If Γ is an elastic graph with elastic constants α, the *extremal length* of c is

$$\text{EL}(c) := \int_{y \in \Gamma} n_c(y)^2 \, dy, \tag{2.5}$$

where we view each edge e of G as having a measure of total mass $\alpha(e)$. This generalizes Definition 1.19; $\text{EL}(c)$ is the ℓ^2 norm of n_c, squared.

- If Γ is a length graph with lengths ℓ, the *weighted length* of c was defined in Equation (1.27); by Lemma 2.10, it is the ℓ^1 norm of n_c.

In Section 3 we will show that each homotopy class has a *taut map* that simultaneously minimizes $n_c(y)$ for all regular values $y \in \Gamma$. Taut maps are automatically minimizers of WR, EL, and ℓ, independent of the structure on Γ. If c is taut, n_c is constant on each edge of Γ and Equation (2.5) reduces to

$$\mathrm{EL}(c) = \sum_{e \in \mathrm{Edge}(G)} n_c(e)^2 \cdot \alpha(e)$$

In addition to these energies, the energies Emb, Dir, and Lip were defined in Section 1. For all of these energies, if we wish to make explicit the dependence of the energy on the geometric structure, we use superscripts for the structure on the domain and subscripts for the structure on the range. Thus we write $\mathrm{WR}_{w_2}^{w_1}$, EL_α^w, ℓ_ℓ^w, $\mathrm{Emb}_{\alpha_2}^{\alpha_1}$, Dir_ℓ^α, or $\mathrm{Lip}_{\ell_2}^{\ell_1}$.

To work with these energies, we give some elementary lemmas, starting with another formula for Dirichlet energy.

DEFINITION 2.6. For G an elastic graph, Γ either a length graph or an elastic graph, and $f: G \to \Gamma$ a PL map, the *filling function* $\mathrm{Fill}_f: \Gamma \to \mathbb{R}$ is defined at regular values of f by

$$\mathrm{Fill}_f(y) := \sum_{x \in f^{-1}(y)} |f'(x)|. \tag{2.7}$$

In particular, Equation (1.15) says that $\mathrm{Emb}(f) = \operatorname{ess\,sup}_y \mathrm{Fill}_f(y)$.

LEMMA 2.8. *For $f: G \to K$ a PL map from an elastic graph to a length graph,*

$$\mathrm{Dir}(f) = \int_{x \in \Gamma} |f'|^2 \, dx = \int_{y \in K} \mathrm{Fill}_f(y) \, dy. \tag{2.9}$$

There is a corresponding formula for ℓ.

LEMMA 2.10. *For $c: W \to K$ a PL map from a weighted graph to a length graph,*

$$\ell(c) = \int_{x \in W} w(x)|c'(x)| \, dx = \int_{y \in K} n_c(y) \, dy. \tag{2.11}$$

Proof of Lemmas 2.8 and 2.10. Change of variables.   □

LEMMA 2.12. *Let $G_1 \xrightarrow{\phi} G_2 \xrightarrow{\psi} \Gamma$ be a sequence of PL maps between graphs, where G_1 and G_2 are elastic graphs and Γ is a length graph or an elastic graph. Then for almost every $z \in \Gamma$,*

$$\mathrm{Fill}_{\psi \circ \phi}(z) = \sum_{y \in \psi^{-1}(z)} |\psi'(y)| \, \mathrm{Fill}_\phi(y) \leqslant \mathrm{Emb}(\phi) \, \mathrm{Fill}_\psi(z).$$

LEMMA 2.13. *Let $W_1 \xrightarrow{\phi} W_2 \xrightarrow{c} \Gamma$ be a sequence of PL maps, with W_1 and W_2 weighted graphs. Then for almost every $z \in \Gamma$,*

$$n_{c \circ \phi}(z) \leqslant \mathrm{WR}(\phi) n_c(z).$$

LEMMA 2.14. *Let $\Gamma \xrightarrow{f} K_1 \xrightarrow{\phi} K_2$ be a sequence of PL maps, with K_1 and K_2 length graphs. Then for almost every $x \in \Gamma$,*

$$\left| (\phi \circ f)'(x) \right| \leqslant |f'(x)| \, \mathrm{Lip}(\phi).$$

Proof of Lemmas 2.12, 2.13, and 2.14. Immediate from the definitions. □

We now turn to submultiplicativity, Equations (1.31) and (1.32). For concreteness, we list the 10 cases individually.

PROPOSITION 2.15. *The energies from (1.26) are submultiplicative, in the following sense. For $i \in \{1, 2, 3\}$, let W_i be marked weighted graphs, G_i be marked elastic graphs, and K_i be marked length graphs. Then, if we are given marked PL maps between these spaces as specified on each line, we have the given inequality.*

$$W_1 \xrightarrow{\phi} W_2 \xrightarrow{\psi} W_3: \qquad \mathrm{WR}(\psi \circ \phi) \leqslant \mathrm{WR}(\phi) \, \mathrm{WR}(\psi) \qquad (2.16)$$

$$W_1 \xrightarrow{\phi} W_2 \xrightarrow{c} G: \qquad \mathrm{EL}(c \circ \phi) \leqslant \mathrm{WR}(\phi)^2 \, \mathrm{EL}(c) \qquad (2.17)$$

$$W_1 \xrightarrow{\phi} W_2 \xrightarrow{f} K: \qquad \ell(f \circ \phi) \leqslant \mathrm{WR}(\phi)\ell(f) \qquad (2.18)$$

$$W \xrightarrow{c} G_1 \xrightarrow{\phi} G_2: \qquad \mathrm{EL}(\phi \circ c) \leqslant \mathrm{EL}(c) \, \mathrm{Emb}(\phi) \qquad (2.19)$$

$$W \xrightarrow{c} G \xrightarrow{f} K: \qquad \ell(f \circ c)^2 \leqslant \mathrm{EL}(c) \, \mathrm{Dir}(f) \qquad (2.20)$$

$$W \xrightarrow{c} K_1 \xrightarrow{\phi} K_2: \qquad \ell(\phi \circ c) \leqslant \ell(c) \, \mathrm{Lip}(\phi) \qquad (2.21)$$

$$G_1 \xrightarrow{\phi} G_2 \xrightarrow{\psi} G_3: \qquad \mathrm{Emb}(\psi \circ \phi) \leqslant \mathrm{Emb}(\phi) \, \mathrm{Emb}(\psi) \qquad (2.22)$$

$$G_1 \xrightarrow{\phi} G_2 \xrightarrow{f} K: \qquad \mathrm{Dir}(f \circ \phi) \leqslant \mathrm{Emb}(\phi) \, \mathrm{Dir}(f) \qquad (2.23)$$

$$G \xrightarrow{f} K_1 \xrightarrow{\phi} K_2: \qquad \mathrm{Dir}(\phi \circ f) \leqslant \mathrm{Dir}(f)\,\mathrm{Lip}(\phi)^2 \tag{2.24}$$

$$K_1 \xrightarrow{\phi} K_2 \xrightarrow{\psi} K_3: \qquad \mathrm{Lip}(\psi \circ \phi) \leqslant \mathrm{Lip}(\phi)\,\mathrm{Lip}(\psi). \tag{2.25}$$

(When there is only one graph of a particular type, we omit the subscript.) Each inequality still holds if we take homotopy classes on both sides.

Proof. Equations (2.16), (2.17), and (2.18) follow from Lemma 2.13 and Equations (1.29), (2.5), and (2.11), respectively. Equations (2.21), (2.24), and (2.25) follow from Lemma 2.14 and Equations (2.11), (1.7), and (1.28), respectively.

For Equation (2.19), we have

$$
\begin{aligned}
\mathrm{EL}(\phi \circ c) &= \int_{z \in G_2} n_{\phi \circ c}(y)^2\, dz \\
&= \int_{z \in G_2} \left(\sum_{\phi(y)=z} n_c(y) \right)^2 dz \\
&\leqslant \int_{z \in G_2} \left(\sum_{\phi(y)=z} n_c(y)^2/|\phi'(y)| \right) \left(\sum_{\phi(y)=z} |\phi'(y)| \right) dz \\
&\leqslant \mathrm{Emb}(\phi) \int_{y \in G_1} n_c(y)^2\, dy \\
&= \mathrm{Emb}(\phi)\,\mathrm{EL}(c),
\end{aligned}
$$

using the definition of EL, the equality $n_{\phi \circ c}(y) = \sum_{\phi(x)=y} n_c(x)$ at regular values, the Cauchy–Schwarz inequality, the definition of $\mathrm{Emb}(\phi)$ and a change of variables between G_2 and G_1, and the definition of EL again.

For Equation (2.20), we have

$$
\begin{aligned}
\ell(f \circ c)^2 &= \left(\int_{z \in K} n_{f \circ c}(z)\, dz \right)^2 \\
&= \left(\int_{y \in G} n_c(y)|f'(y)|\, dy \right)^2 \\
&\leqslant \left(\int_{y \in G} n_c(y)^2\, dy \right) \left(\int_{y \in G} |f'(y)|^2\, dy \right) \\
&= \mathrm{EL}(c)\,\mathrm{Dir}(f),
\end{aligned}
$$

using Lemma 2.10, a change of variables from K to G, the Cauchy–Schwarz inequality, and the definitions of EL and Dir.

Equation (2.22) follows from Lemma 2.12.

For Equation (2.23), we have

$$
\begin{aligned}
\mathrm{Dir}(f \circ \phi) &= \int_{z \in K} \mathrm{Fill}_{f \circ \phi}(z)\, dz \\
&\leqslant \int_{z \in K} \mathrm{Fill}_f(z) \Big(\max_{f(y)=z} \mathrm{Fill}_\phi(y) \Big)\, dz \\
&\leqslant \int_{z \in K} \mathrm{Fill}_f(z)\, \mathrm{Emb}(\phi)\, dz \\
&= \mathrm{Emb}(\phi)\, \mathrm{Dir}(f),
\end{aligned}
$$

using Lemmas 2.8, 2.12, the definition of $\mathrm{Emb}(\phi)$, and Lemma 2.8 again.

For each of these equations, to replace concrete functions by homotopy classes, take representatives f, g of the two homotopy classes whose energy is within a factor of ε of the optimal value, in the sense that $E_q^p[f] \leqslant E_q^p(f) \leqslant E_q^p[f](1+\varepsilon)$, and similarly for g. Then

$$
E_r^p[g \circ f] \leqslant E_r^p(g \circ f) \leqslant E_q^p(f) E_r^q(g) \leqslant E_q^p[f] E_r^q[g](1 + \varepsilon)^2.
$$

Since ε can be chosen as small as desired, we are done. $\qquad\square$

We now have one direction of Theorem 1.

COROLLARY 2.26. *For any homotopy class* $[\phi]\colon G_1 \to G_2$ *of maps between marked elastic graphs,* $\mathrm{SF}_{\mathrm{Dir}}[\phi] \leqslant \mathrm{Emb}[\phi]$ *and* $\mathrm{SF}_{\mathrm{EL}}[\phi] \leqslant \mathrm{Emb}[\phi]$.

Proof. For any marked length graph K and homotopy class $[f]\colon G_2 \to K$, by the homotopy version of Equation (2.23) we have

$$
\frac{\mathrm{Dir}[f \circ \phi]}{\mathrm{Dir}[f]} \leqslant \frac{\mathrm{Dir}[f]\, \mathrm{Emb}[\phi]}{\mathrm{Dir}[f]} = \mathrm{Emb}[\phi].
$$

Since K and $[f]$ were arbitrary, it follows that $\mathrm{SF}_{\mathrm{Dir}}[\phi] \leqslant \mathrm{Emb}[\phi]$. The argument that $\mathrm{SF}_{\mathrm{EL}}[\phi] \leqslant \mathrm{Emb}[\phi]$ is exactly parallel. $\qquad\square$

3. Reduced and taut maps

We next turn to notions of efficiency of maps between graphs. There is a weak notion depending on no extra structure ('reduced'), and then a more powerful notion that depends on a weight structure ('taut').

3.1. Reduced maps. We work by analogy with the standard notion of a reduced (cyclic) word in a group.

DEFINITION 3.1. A map $f\colon \Gamma_1 \to \Gamma_2$ between marked graphs is *edge-reduced* if, on each edge e of Γ_1, f is either constant or has a perturbation that is locally injective.

DEFINITION 3.2. For a graph Γ and point $x \in \Gamma$, a *direction* d at x is a germ of a PL map from $\mathbb{R}_{\geqslant 0}$ to Γ starting at x, considered up to PL reparametrization. Explicitly, there is a *zero* direction, the germ of a constant map. Points on edges of Γ have two nonzero directions, and vertices have as many nonzero directions as their valence. If $f\colon \Gamma_1 \to \Gamma_2$ is a PL map and d is a direction at $x \in \Gamma_1$, then $f(d)$ is a direction at $f(x)$.

DEFINITION 3.3. A locally injective PL map $f\colon \Gamma_1 \to \Gamma_2$ between marked graphs is *reduced* if, at each unmarked vertex v of Γ_1, there are directions d_1 and d_2 at v so that $f(d_1)$ and $f(d_2)$ are distinct and nonzero.

More generally, pick $y \in \Gamma_2$ and let $Z \subset \Gamma_1$ be a connected component of $f^{-1}(y)$. A *direction* from Z is a point $x \in Z$ and a direction d from x that points out of Z (so that $f(d)$ is nonzero). Then Z is a *dead end* for f if Z has no marked point and there is exactly one direction in

$$D(Z) := \{\, f(d) \mid d \text{ a direction from } Z \,\}.$$

(If $D(Z) = \varnothing$, then Z is an entire connected component of Γ_1. We do not count this as a dead end.) We say that f is *reduced* if it has no dead ends. If f is not reduced, there is a natural *reduction* operation at a dead end Z, where we modify f by pulling the image of Z in the direction $D(Z)$ until it hits a vertex of Γ_2 or the boundary of a domain of linearity.

PROPOSITION 3.4. *If $f\colon \Gamma_1 \to \Gamma_2$ is any map between marked graphs, then repeated reduction makes f into a reduced map. In particular, there is a reduced map in every homotopy class.*

Proof. Proceed by induction on the number of linear segments of f with nonzero derivative. This is reduced by each reduction operation. □

PROPOSITION 3.5. *For $p, q \in \{1, 2, \infty\}$ with $1 \leqslant p \leqslant q \leqslant \infty$, reduction does not increase E_q^p, and strictly decreases E_q^p if $p < q$.*

Proof. Clear from the definitions. □

As a result of Propositions 3.4 and 3.5, when looking for energy minimizers we can restrict our attention to reduced maps.

3.2. Taut maps and flows.   We now add some more structure, and correspondingly get more restrictive conditions on energy-minimizing maps. We will consider maps from weighted graphs, and in particular minimize E_q^1 for any q.

DEFINITION 3.6. Let $c \colon W \to \Gamma$ be a PL map from a marked weighted graph W to a marked graph Γ. We defined the multiplicity $n_c \colon \Gamma \to \mathbb{R}_{\geqslant 0}$ in Definition 2.4. For a homotopy class $[c]$ and y in the interior of an edge of Γ set

$$n_{[c]}(y) := \inf_{d \in [c]} n_d(y).$$

The infimum is taken over maps d for which $n_d(y)$ is defined. Then it is easy to see that $n_{[c]}(y)$ depends only on the edge containing y. We say that c is *taut* if $n_c = n_{[c]}$ almost everywhere. We say that c is *locally taut* if for every $y \in \Gamma$ and regular neighborhood N of y, the quantity n_c cannot be reduced by homotopy of c on $c^{-1}(N)$.

It will require work to show that taut maps exist, but if they do exist they have good properties.

LEMMA 3.7. *A taut map from a positive weighted graph is reduced.*

Proof. Reduction strictly reduces n_c.   □

PROPOSITION 3.8. *Let $c \colon W \to \Gamma$ be a taut map from a marked weighted graph W to a marked graph Γ. If Γ is weighted, then $\mathrm{WR}(c)$ is minimal in $[c]$. If Γ has an elastic structure, then $\mathrm{EL}(c)$ is minimal in $[c]$. If Γ has a length structure, then $\ell(c)$ is minimal in $[c]$.*

Proof. Clear from the definitions, since the energies are monotonic in n_c. When Γ is a length graph, we use Lemma 2.10.   □

See also Proposition 3.18.

EXAMPLE 3.9. Let W be a marked weighted graph with two marked vertices s and t. Let I be the interval $[-1, 1]$ with the endpoints marked, and let $f \colon W \to I$ be a map with $f(s) = -1$ and $f(t) = 1$. Then f is taut if and only if each edge is mapped monotonically and, for each regular value $y \in I$, the set of edges

containing $f^{-1}(y)$ is a *minimal cut-set* for W, a set of edges of W that separates s from t and has minimal weight among all such sets.

The max-flow/min-cut theorem says that in the context of Example 3.9 the total weight of a minimal cut-set for W is equal to the maximum flow from s to t through the edges of W. We will show that taut maps exist, and that they satisfy a generalization of the max-flow/min-cut theorem; see Theorem 3 and Corollary 3.32. To give the strongest statement, we make some definitions and preliminary lemmas first.

LEMMA 3.10. *A marked weighted multicurve* $c\colon C \to \Gamma$ *on a marked graph* Γ, *with weights that are positive and constant on the components of* C, *is taut if and only if it is reduced.*

Proof. One direction is Lemma 3.7. For the other direction, if c is reduced and nonconstant, perturb it to a locally injective map and use the standard fact that, given a nontrivial free homotopy class of maps from S^1 to Γ, a locally injective representative is unique up to reparametrization of the domain (which does not change n_c). The same is true for homotopy classes of maps from I to Γ relative to the endpoints. $\square$

DEFINITION 3.11. If $c\colon W_1 \to W_2$ is a PL map between weight graphs, we say that W_2 *carries* (W_1, c) if $\mathrm{WR}(c) \leqslant 1$. If W_2 carries (W_1, c), then we say that a point y of W_2 is *saturated* if $n_c(y) = w(y)$. Similarly, a subset of a weighted graph is saturated if almost every point in it is saturated.

One notion of a 'flow' on a weighted graph W is a taut weighted multicurve carried by W. These multicurves are a little awkward to work with directly. As such, we will also define and work with a more general type of flow from *train tracks*.

DEFINITION 3.12. A sequence of nonnegative numbers $(x_i)_{i=1}^k$ is said to satisfy the *triangle inequalities* if, for each i, x_i is no larger than the sum of the remaining numbers:

$$x_i \leqslant \sum_{\substack{1 \leqslant j \leqslant k \\ i \neq j}} x_j. \tag{3.13}$$

This implies that $k \neq 1$, and if $k = 2$ then $x_1 = x_2$. If $k \geqslant 3$ and one of these inequalities is an equality, then there is exactly one i so that x_i is equal to the sum of the remainder. That x_i is said to *dominate* the rest.

DEFINITION 3.14. A *train-track structure* τ on a graph Γ is, for each point x of Γ, a partition of the nonzero directions from x into equivalence classes, called the *gates* at x, with at least two gates at each unmarked point. (This only gives additional structure at the vertices of Γ.) A *train track* is a graph with a train-track structure.

If the graph Γ of a train track is weighted, the *weight* of a gate is the sum of the weights of the edges corresponding to the directions that make it up. A set of weights on a train track is *balanced* if, at each unmarked vertex, the weights of the gates satisfy the triangle inequalities. A *balanced train track* is a train track together with balanced weights. If the weight of one gate g at a vertex of a balanced train track dominates the others, we can *smooth* the vertex, changing the train-track structure so that there are only two gates, one with the directions from g and one with all the other directions.

A graph Γ with no unmarked univalent ends has a *discrete* train-track structure δ_Γ, in which two different directions are never equivalent. By default we use the discrete train-track structure on a graph. A (marked) weighted graph for which the discrete train-track structure satisfies the triangle inequalities (at unmarked vertices) is said to be *balanced*.

If $f\colon (\Gamma_1, \tau_1) \to (\Gamma_2, \tau_2)$ is a map between train tracks that is locally injective on the edges of Γ_1, we say that f is a *train-track map* if, for each vertex v of Γ_1 and directions d_1 and d_2 at v,

$$\big(d_1 \sim_{\tau_1} d_2\big) \Longleftrightarrow \big(f(d_1) \sim_{\tau_2} f(d_2)\big). \tag{3.15}$$

More generally, we say that f is a train-track map if it has arbitrarily small perturbations f_ε so that f_ε is locally injective on the edges of Γ_1 and satisfies Equation (3.15). In particular, a multicurve $c\colon C \to \Gamma$ is a train-track map if and only if at each point it passes through the incoming and outgoing directions are in different gates.

If $f\colon \Gamma_1 \to \Gamma_2$ is a locally injective, reduced map, the *train track of* f is the unique train-track structure $\tau(f)$ on Γ_1 so that f is a train-track map with respect to the train-track structures $\tau(f)$ on Γ_1 and δ_{Γ_2} on Γ_2. Concretely, we have

$$\big(d_1 \sim_{\tau(f)} d_2\big) \Longleftrightarrow \big(f(d_1) = f(d_2)\big).$$

REMARK 3.16. Train tracks were first introduced in the context of surfaces and the theory of pseudo-Anosov diffeomorphisms [**PH92**], and a slight different notion of train-track maps between graphs is used to study automorphisms of free groups [**BH92**].

LEMMA 3.17. *A composition of train-track maps is a train-track map.*

Proof. This is obvious if both maps are locally injective. In general, given train-track maps f and g with locally injective train-track perturbations f_ε and g_ε, the composition $g_\varepsilon \circ f_\varepsilon$ is a locally injective train-track perturbation of $g \circ f$. $\qquad\square$

PROPOSITION 3.18. *Let $f : W \to \Gamma$ be a taut map from a marked weighted graph. Then $n_f(\cdot)$ gives Γ the structure of a balanced weighted graph.*

Proof. If n_f (as a function on the edges of Γ) is not balanced at an unmarked vertex v, with one incident edge e dominating the others, we may homotop f to move the image of $f^{-1}(v)$ along e. This decreases n_f, contradicting the assumption that f was taut. $\qquad\square$

We can now state the main goal of this section.

THEOREM 3. *Let $f : W \to \Gamma$ be a PL map from a marked weighted graph to a marked graph. Then there is a taut map in $[f]$. Furthermore, the following conditions are equivalent.*

(1) *The map f is taut.*

(2) *The map f is locally taut.*

(3) *The graph W carries a marked balanced train track $t : T \to W$, so that $f \circ t$ is a train-track map (with respect to the discrete train-track structure on Γ) and $n_{f \circ t} = n_f$. We may choose (T, t) so that T is a subgraph of W.*

(4) *The graph W carries a marked weighted multicurve $c : C \to W$ so that $f \circ c$ is taut and $n_{f \circ c} = n_f$.*

Furthermore, if these conditions are satisfied, t in (3) and c in (4) saturate every edge of W on which f is not constant, and c can be chosen to factor through t: there is a multicurve $c' : C \to T$ so that $c = t \circ c'$.

As a preliminary step toward Theorem 3, we relate train tracks and multicurves.

PROPOSITION 3.19. *Any balanced train track $T = (\Gamma, w, \tau)$ carries a marked weighted multicurve (C, c) that saturates T and so that $c : C \to T$ is a train-track map. We can choose (C, c) so that each component of C runs over each edge of Γ at most twice.*

Proof. Let $T' \subset T$ be the subtrain track of edges of nonzero weight. Let T'' be T' modified by smoothing all vertices in which one gate dominates the others. Let

$\text{Yard}(T'') \subset \text{Vert}(T'')$ be the set of marked vertices of T'' with at least three gates. We will proceed by induction on $|\text{Edge}(T'')| + |\text{Yard}(T'')|$.

In the base case, T' and T'' are empty.

Otherwise, choose any oriented edge $\vec{e}_0$ of T'', and find a path (forward and backward) from $\vec{e}_0$ within T'', always making turns between different gates. Since there are at least two gates at each unmarked vertex, we can always find a successor edge for $\vec{e}_i$, unless we hit a marked vertex. Since there are only finitely many oriented edges in T, we will eventually either find a path of edges between marked points or see a repeat and find a cyclic loop of edges.

Consider the marked curve (C_1, c_1) that runs over the cycle or path. For $\varepsilon > 0$, let $w_\varepsilon(e) := w(e) - \varepsilon n_{c_1}(e)$. Then for ε sufficiently small, w_ε gives a balanced train-track structure on T'':

- On an edge e, $w(e) > 0$ by construction of T' so $w_\varepsilon(e) \geqslant 0$; and

- At a vertex v, the construction of T'' ensures that the triangle inequalities at v continue to hold in w_ε.

Let ε_1 be the maximum value of ε so that $(\Gamma, w_{\varepsilon_1}, \tau)$ is a balanced train track. Let $w_1 = w_{\varepsilon_1}$ and $T_1 = (\Gamma, w_1, \tau)$. As before, we construct derived train tracks T_1' by deleting edges of weight 0 from T_1 and T_1'' by smoothing dominating vertices in T_1'. By the choice of ε_1, there is either:

- an edge e of T'' with $w_1(e) = 0$ but $w(e) \neq 0$, so $|\text{Edge}(T_1'')| < |\text{Edge}(T'')|$; or

- a vertex v of Γ with a gate g so that $w_1(g)$ dominates the other gates but $w(g)$ does not, so $|\text{Yard}(T_1'')| < |\text{Yard}(T'')|$.

In either case, by induction T_1 carries a marked weighted multicurve (C_2, c_2) that saturates T_1. Then $(C_1 \sqcup C_2, c_1 \sqcup c_2, \varepsilon_1 \sqcup w_2)$ is the desired multicurve from the statement.

If, in the construction above, we make a cyclic loop as soon as we see a repeated oriented edge, the components of C run over each (unoriented) edge at most twice. $\quad\square$

LEMMA 3.20. *If $f : W \to \Gamma$ is a map from a marked weighted graph to a marked graph and (C, c) is a marked weighted multicurve carried by W so that $f \circ c$ is taut and $n_{f \circ c} = n_f$, then f is taut.*

Proof. If $g : W \to \Gamma$ is any other map in $[f]$, then

$$n_g \geqslant n_{g \circ c} \geqslant n_{[f \circ c]} = n_{f \circ c} = n_f,$$

using the fact that W carries c and Lemma 2.13, the definition of $n_{[f \circ c]}$ as an infimum over the homotopy class, and the hypotheses. □

COROLLARY 3.21. *A train-track map from a marked balanced train track is taut.*

Proof. Let $f : T \to \Gamma$ be the train-track map. Let $c : C \to T$ be the multicurve from Proposition 3.19. Then, by Lemma 3.17, $f \circ c$ is taut and, since c saturates T, $n_{f \circ c} = n_f$. Thus by Lemma 3.20, f is taut. □

3.3. Local models. To prove Theorem 3, we first analyze the situation locally in a regular neighborhood of a singular value. This reduces to studying maps from a graph with k marked points to a k-leg star graph.

Let Star_k be the star graph with k legs, with a central vertex s_*, marked endpoints $s_1, \ldots, s_k$, and edges $[s_i, s_*]$. A *k-marked* graph is a graph with k marked vertices $(v_i)_{i=1}^k$ (in order). There is a canonical homotopy class of marked maps from a k-marked graph to Star_k, taking v_i to s_i. We prove an analogue of Theorem 3 in this context.

PROPOSITION 3.22. *Let W be a k-marked weighted graph. Then there is a taut map in the canonical homotopy class of maps to Star_k. Furthermore, the following conditions are equivalent.*

(1) *The map f is taut.*

(2) *The graph W carries a marked balanced train track (T, t) so that $f \circ t$ is a train-track map and $n_{f \circ t} = n_f$.*

(3) *The graph W carries a marked weighted multicurve (C, c) so that $f \circ c$ is taut and $n_{f \circ c} = n_f$.*

DEFINITION 3.23. A *cut* S of a graph Γ is a partition of the vertices of the graph into two disjoint subsets: $\mathrm{Vert}(\Gamma) = S \sqcup \overline{S}$. The corresponding *cut-set* $c(S) = c(\overline{S})$ is the set of edges that connect S to $\overline{S}$. If Γ has weights w, the *weight* of the cut is $w(S) := \sum_{e \in c(S)} w(e)$.

Two cuts S_1 and S_2 are *nested* if they are disjoint or one is contained in the other: $S_1 \cap S_2 = \varnothing$, $S_1 \subset S_2$, or $S_2 \subset S_1$. The definition is unchanged if we replace an S_i with $\overline{S_i}$. A set of cuts is nested if each pair is nested.

Let W be a k-marked weighted graph. A *v_i-cut* is a subset $S_i \subset \mathrm{Vert}(G)$ so that $S_i \cap \{v_1, \ldots, v_k\} = \{v_i\}$. A *vertex cut* is a v_i-cut for some i. A *minimal v_i-cut* is one with minimal weight. Let $\mathrm{mincut}_i(w)$ be the weight of a minimal v_i-cut.

LEMMA 3.24. *If S_1 and S_2 are two cuts on the same weighted graph, then $w(S_1) + w(S_2) \geqslant w(S_1 \cap S_2) + w(S_1 \cup S_2)$.*

Proof. This is the standard submodular property of cuts, and is easy to prove. ☐

LEMMA 3.25. *If W is a k-marked weighted graph and S_i and S_i' are two minimal v_i-cuts, then $S_i \cap S_i'$ and $S_i \cup S_i'$ are also minimal v_i-cuts.*

Proof. The sets $S_i \cap S_i'$ and $S_i \cup S_i'$ are v_i-cuts, so by minimality they both have weight at least as large as $\mathrm{mincut}_i(w) = w(S_i) = w(S_i')$. Lemma 3.24 gives the other inequality. ☐

LEMMA 3.26. *If W is a k-marked weighted graph, S_i is a minimal v_i-cut, and S_j is a minimal v_j cut for $j \neq i$, then $S_i \backslash S_j$ is also a minimal v_i-cut and $S_j \backslash S_i$ is also a minimal v_j-cut.*

Proof. $S_i \backslash S_j$ is a v_i-cut and $S_j \backslash S_i$ is a v_j-cut. By minimality, their weights are at least as large as $\mathrm{mincut}_i(w)$ and $\mathrm{mincut}_j(w)$, respectively. Applying Lemma 3.24 to S_i and $\overline{S_j}$ gives the other inequalities. ☐

DEFINITION 3.27. If W is a k-marked weighted graph, we say that an edge of W is *slack* if it has nonzero weight and is not contained in any minimal vertex cut.

See Figure 2 for the next two lemmas.

LEMMA 3.28. *If $W = (\Gamma, w)$ is a weighted graph with k marked vertices, then there is a set of weights $w_0 \leqslant w$ on Γ so that $W_0 = (\Gamma, w_0)$ has no slack edges and so that for all i, $\mathrm{mincut}_i(w) = \mathrm{mincut}_i(w_0)$.*

Proof. We proceed by induction on the number of slack edges. If there is one, pick any slack edge e_0 of W. For $0 \leqslant k \leqslant w(e_0)$, define a modified set of weights by

$$w\{e_0 \mapsto k\}(e) := \begin{cases} k & e = e_0, \\ w(e) & \text{otherwise.} \end{cases}$$

Let k_0 be minimal value so that, for all i, $\mathrm{mincut}_i(w\{e_0 \mapsto k_0\}) = \mathrm{mincut}_i(w)$. Then $k < w(e_0)$ and e is not slack with respect to $w\{e_0 \mapsto k_0\}$. By induction we can find weights $w_0 \leqslant w\{e_0 \mapsto k_0\} \leqslant w$ on Γ so that (Γ, w_0) has no slack edges. ☐

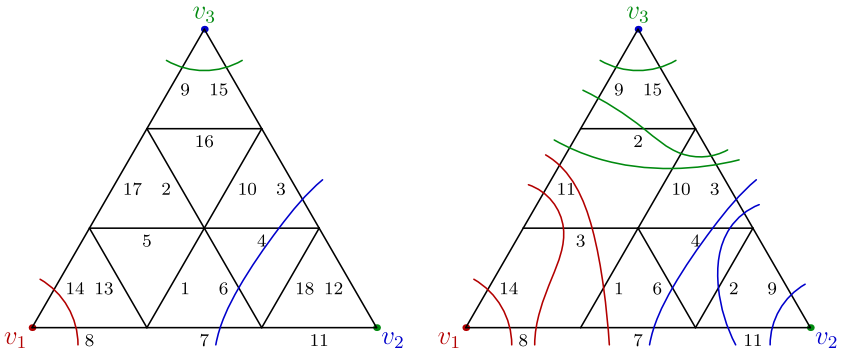The weights w_0 in Lemma 3.28 are usually not unique.

Figure 2. Finding minimum cuts and maximum flows near a vertex. Left: a weighted graph W with three marked vertices, with the weights indicated. Minimal cuts that separate the vertices are marked; these are used to construct a taut map from W to Star_3. Right: a weighted graph W_0 carried by W, with no slack edges and the same mincut_i for each i. The minimal cuts on W have been extended to a complete nested family of minimal cuts. These are used to construct a train track carried by W.

LEMMA 3.29. *Let $W = (\Gamma, w)$ be a k-marked weighted graph with no slack edges and let $\mathcal{S} \subset \mathcal{P}(\mathcal{P}(\mathrm{Vert}(W)))$ be a nested set of minimal vertex cuts. Then there is a nested set of minimal vertex cuts $\mathcal{T} \supset \mathcal{S}$ so that every edge of W with nonzero weight is contained in $c(S)$ for some $S \in \mathcal{T}$.*

Here $\mathcal{P}(X)$ is the power set of X.

Proof. We proceed by induction on the size of $\mathrm{Edge}(W) \backslash \bigcup_{S \in \mathcal{S}} c(S)$. If there is at least one edge e_1 in this set with nonzero weight, it is contained in some minimal vertex cut-set $c(S_1)$ since W has no slack edges. Then, by repeatedly applying Lemmas 3.25 and 3.26, we can find another minimal vertex cut S_1' so that $e \in c(S_1')$ and S_1' is nested with respect to all $S \in \mathcal{S}$. By induction, $\mathcal{S} \cup \{S_1'\}$ can be completed to the desired set $\mathcal{T}$. □

Proof of Proposition 3.22. We first prove the existence of a taut map. For each i, pick a minimal v_i-cut S_i, and let $S_i' = S_i \backslash \bigcup_{j \neq i} S_j$. By Lemma 3.26, $\{S_i'\}$ is a nested collection of minimal v_i-cuts. Define $f : W \to \mathrm{Star}_k$ by mapping all vertices in S_i' to s_i, all vertices not in any S_i' to s_*, and all edges to reduced paths connecting their endpoints. (Thus f is constant on any edge not in any $c(S_i')$.) Then f is taut since the S_i' are minimal.

Proposition 3.19 and Lemma 3.10 tell us that (2) implies (3), and Lemma 3.20 tells us that (3) implies (1), so it remains to prove that (1) implies (2).

Suppose that f is taut. For each regular value $y \in [s_*, s_i] \subset \text{Star}_k$, consider the v_i-cut $S_y = \{v \in \text{Vert}(\Gamma) \mid f(v) \in [s_i, y]\}$. Then S_y is a minimal vertex cut, since f is taut. Furthermore, if z is another regular value of f on any edge of Star_k, then S_y and S_z are nested. Let $\mathcal{S}$ be the set of all S_y for all regular y, together with the singleton endpoints $\{v_i\}_{i=1}^k$. This is a nested set of minimal cuts on W.

Use Lemma 3.28 to find weights $w_0 \leqslant w$ so that (Γ, w_0) has no slack edges. Let $\Gamma_0 \subset \Gamma$ consist of the edges e with $w_0(e) \neq 0$, let $W_0 = (\Gamma_0, w_0)$, and let $t \colon W_0 \hookrightarrow W$ be the inclusion. By Lemma 3.29, $\mathcal{S}$ can be completed to a nested system of cuts $\mathcal{T}$ on W_0 so that every edge is in at least one cut-set. For one example of W_0 and $\mathcal{T}$, see the right of Figure 2.

For each i, let the distinct v_i-cuts in $\mathcal{T}$ be

$$\{v_i\} = S_{i,0} \subsetneq S_{i,1} \subsetneq \cdots \subsetneq S_{i,n_i}.$$

For each i and for $0 \leqslant j \leqslant n_i$, pick distinct points $x_{i,j} \in [s_i, s_*] \subset \text{Star}_k$ with $x_{i,0} = s_i$ and with the $x_{i,j}$ in order. Then define a map $g \colon W_0 \to \text{Star}_k$ on vertices by

$$g(v) = \begin{cases} s_i & v = v_i, \\ x_{i,j} & v \in S_{i,j} \backslash S_{i,j-1}, \\ s_* & v \in \text{Vert}(W_0) \backslash \bigcup_{S \in \mathcal{T}} S. \end{cases}$$

(These cases are exclusive since $\mathcal{T}$ is nested.) On an edge e of W_0, define $g(e)$ to be a reduced path connecting its endpoints. Since $\mathcal{T}$ is a complete set of cuts, g is not constant on any edge. Let $\tau_0 = \tau(g)$. Then (W_0, τ_0) is a balanced train track: the triangle inequalities follow since the $S_{i,j}$ are minimal cuts.

By an appropriate choice of the $x_{i,j}$, the train-track map g can be made to be an arbitrarily small perturbation of $f \circ t$, so $f \circ t$ is also a train-track map. If e is an edge on which the original map f is not constant, then e is not slack and $w(e) = w_0(e)$, so $n_{f \circ t} = n_f$ as desired. $\qquad \square$

3.4. General flows. We now use these local models to prove Theorem 3.

DEFINITION 3.30. Let $f \colon W \to \Gamma$ be a PL map from a marked weighted graph W to a marked graph Γ. Pick $y \in \Gamma$ and let $N \subset \Gamma$ be a closed regular neighborhood of y. Then the *local model for f at y* is the map

$$f^{-1}(N)/{\sim} \longrightarrow N,$$

where:

- N is considered as a marked graph (equivalent to Star_k) with marked points at ∂N;

- $\sim$ is the equivalence relation that identifies two points in $f^{-1}(\partial N)$ if they map to the same point in ∂N; and

- $f^{-1}(N)/\sim$ is considered as a k-marked weighted graph, with weights inherited from W.

LEMMA 3.31. *Let $[f]\colon W \to \Gamma$ be a homotopy class of maps from a marked weighted graph W to a marked graph Γ. Then there is a locally taut map in $[f]$.*

Proof. Suppose f is any PL representative for its homotopy class. We can modify f to send vertices to vertices without increasing n_f, as follows. For each edge e of Γ, look at the division of e into intervals according to the value of n_f. Pick one of these intervals on which n_f is minimal (among the values of n_f that appear), and spread out this interval by a homotopy until it covers e. This gives us an initial map f_0.

If f_0 is not locally taut, then there is some vertex v of Γ so that the local model of f_0 at v is not taut. By Proposition 3.22, there is a taut map in the homotopy class of the local model. Let f_0' be f_0 with the map replaced by its taut model near v. There is at least one point x near v, regular for both maps, with $n_{f_0'}(x) < n_{f_0}(x)$. Homotop f_0' as above, spreading out segments of minimal multiplicity, to construct a map f_1 that sends vertices to vertices with $n_{f_1} \leqslant n_{f_0}$ everywhere and $n_{f_1}(x) < n_{f_0}(x)$.

If f_1 is not locally taut, repeat the process, with f_1 in place of f_0. Our initial representative f_0 gives an upper bound on $n_{f_i}(e)$ for all e. There are only finitely many linear combinations of the nonzero weights of edges of W to get a value less than this bound. At each step n_f strictly decreases on at least one edge, so the process terminates in finite time. □

Proof of Theorem 3. We first prove the equivalence of the four conditions on taut maps. (1) implies (2) is obvious.

To see that (2) implies (3), suppose f is locally taut. Then for each singular value $y \in \Gamma$, the local model for f at y carries a balanced train track compatible with f by Proposition 3.22. Define a balanced train track T and map $t\colon T \to W$ by assembling these local models following the pattern of W, leaving W unchanged outside of the local models. Then (T, t) is the desired train track carried by W, with T a subgraph of W.

Proposition 3.19 shows that (3) implies (4).

Lemma 3.20 shows that (4) implies (1).

Now if $[f]$ is any homotopy class, by Lemma 3.31 there is a locally taut element of $[f]$, which by the above equivalences is also globally taut.

The statements about saturation follow immediately, and the multicurve c factors through t.   □

3.5. Connection to max-flow/min-cut.

For some intuition, we can connect Theorem 3 and Proposition 3.22 to a statement that looks more like the classical max-flow/min-cut problem (Example 3.9). For simplicity, we restrict to the local setting of Proposition 3.22. Given a graph Γ with k marked points $\{v_i\}_{i=1}^k$, define a *flow* between v_i and v_j to be a marked weighted multicurve (C, c) on Γ, with each component a marked interval with one endpoint mapping to v_i and the other mapping to v_j. Such a flow has multiplicities n_c on each edge as usual, as well as a total weight $w(c)$, the sum of the weights of all components of C.

COROLLARY 3.32. *Let W be a weighted graph with k marked points $\{v_i\}_{i=1}^k$. Then we can find:*

- *for $1 \leqslant i \leqslant k$, a v_i-cut S_i; and*

- *for $1 \leqslant i < j \leqslant k$, a flow $c_{ij} = c_{ji}$ between v_i and v_j*

so that:

- *the collection of all flows c_{ij} is carried by W: for each edge e of Γ;*

$$\sum_{i<j} n_{c_{ij}}(e) \leqslant w(e);\qquad(3.33)$$

- *the total flow into a vertex equals the weight of the cut: for each i;*

$$\sum_{j\neq i} w(c_{ij}) = w(S_i).\qquad(3.34)$$

In this situation, for each i, $w(S_i)$ is minimal and $\sum_{j\neq i} w(c_{ij})$ is maximal.

Proof. Let f be a taut map in the canonical homotopy class of maps from W to Star_k, as given by Proposition 3.22. By that proposition, W carries a marked weighted multicurve (C, c) so that $f \circ c$ is taut and $n_{f\circ c} = n_f$. The only nontrivial homotopy classes of marked multicurves on Star_k are given by paths connecting s_i to s_j for some $i \neq j$. Let the flow c_{ij} be the flow given by those components of (C, c) that connect v_i to v_j. Then Equation (3.33) says that W carries (C, c).
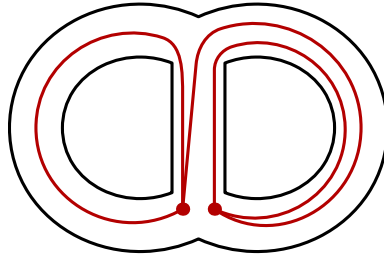
Figure 3. A map from an eyeglass graph to a theta graph that is reduced but not strongly reduced.

On the other hand, for each i, let y_i be a regular value on $[s_*, s_i]$, and let S_i be S_{y_i} as defined in the proof of Proposition 3.22. Then the equality $n_{f \circ c}(y) = n_f(y)$ is equivalent to Equation (3.34). Minimality of $w(S_i)$ and maximality of $\sum_{j \neq i} w(c_{ij})$ both follow from tautness, and indeed are easy to deduce from Equation (3.34). □

3.6. Strongly reduced maps. For later finiteness statements, we introduce a stronger notion of 'reduced' for maps between general marked graphs without a weight structure.

DEFINITION 3.35. A map $f \colon \Gamma_1 \to \Gamma_2$ between marked graphs is *strongly reduced* if it is taut for some choice of positive weights on Γ_1. See Figure 3 for a nonexample.

LEMMA 3.36. *Let $f \colon \Gamma_1 \to \Gamma_2$ be an edge-reduced map. Then f is strongly reduced if and only if for each edge e of Γ_1, there is a reduced curve (C_e, c_e) on Γ_1 that runs over e and so that $f \circ c$ is also reduced. Furthermore, the curves can be chosen to run over each edge of Γ_1 at most twice.*

Proof. If f is strongly reduced, then let $w \in \mathcal{W}^+(\Gamma_1)$ be a set of weights for which f is taut. Then the connected components of the multicurve from condition (4) of Theorem 3 is the desired curve for each edge, which by Proposition 3.19 may be chosen to run over each edge at most twice.

Conversely, if there are curves (C_e, c_e) as in the statement, let $W = \bigsqcup_e C_e$ be their union with the evident map $c \colon W \to \Gamma_1$ and with constant weight $\mathbf{1}$ on each component. Then f is taut with respect to $n_c(\mathbf{1})$. □

DEFINITION 3.37. Given an edge-reduced PL map $f \colon \Gamma_1 \to \Gamma_2$ between marked graphs, the *combinatorial type* of f consists of the following discrete data.

(1) For each vertex v of Γ_1, record whether $f(v)$ is on a vertex or in the interior of an edge of Γ_2, and which vertex or edge it lies on.

(2) For each oriented edge $\vec{e}$ of Γ_1, record the *edge-path* of $f(\vec{e})$, the reduced sequence of oriented edges of Γ_2 that $f(\vec{e})$ travels over. This edge-path may start and/or end with a partial edge, depending on whether the endpoints of $\vec{e}$ map to vertices or edges. There are degenerate cases when $\vec{e}$ maps to a single vertex or stays within the interior of a single edge.

The combinatorial type of f determines the homotopy class of $[f]$.

PROPOSITION 3.38. *Let $[f] \colon \Gamma_1 \to \Gamma_2$ be a homotopy class of maps between marked graphs. Then there are only finitely many combinatorial types of strongly reduced maps in $[f]$.*

Proof. Let $g \in [f]$ be strongly reduced, and let $\vec{e}$ be an oriented edge of Γ_1. We must show that there is a finite set of possible edge-paths for $f(\vec{e})$. For this, let (C, c) be the curve running over e as in Lemma 3.36. Since (C, c) runs over each edge of Γ_1 at most twice, there are only finitely many possibilities for it. Furthermore, $[f \circ c]$ is determined by $[f]$ and $[c]$, so there are only finitely many edge-paths in Γ_2 arising from $[f \circ c]$. The edge-path of $f(\vec{e})$ must be a subpath of the edge-path of $f \circ c$, and again there are only finitely many possibilities. $\qquad\square$

REMARK 3.39. Proposition 3.38 is false if we look at reduced maps rather than strongly reduced maps. (The map f can be 'spun' around a taut cycle of Γ_1, as in Figure 3.)

3.7. Restricting the domain and range.

LEMMA 3.40. *Let $f \colon W \to \Gamma$ be a map from a marked weighted graph to a marked graph. Let $W_0 \subset W$ be the closure of the subset of W where $w(x) \neq 0$. Then f is taut if and only if the restriction $f|_{W_0} \colon W_0 \to \Gamma$ is taut.*

Proof. Segments of weight 0 do not effect n_f. $\qquad\square$

DEFINITION 3.41. A homotopy class $[f] \colon \Gamma_1 \to \Gamma_2$ of marked maps is *essentially surjective* if every element of $[f]$ is surjective.

REMARK 3.42. If Γ_1 and Γ_2 are connected with no 1-valent vertices and $[\phi]\colon \Gamma_1 \to \Gamma_2$ is π_1-surjective, then it is essentially surjective.

If f is not (essentially) surjective, then tautness of f is equivalent to tautness of the map with restricted range.

LEMMA 3.43. *Let $f\colon W \to \Gamma$ be a map from a marked weighted graph to a marked graph. Suppose the image of f is contained in a subgraph $\Gamma_0 \subset \Gamma$. Let Γ^* be Γ with all edges not in Γ_0 collapsed, with $\kappa\colon \Gamma \to \Gamma^*$ the collapsing map. Then the following are equivalent:*

(1) *f is taut;*

(2) *the map with restricted range $f\colon W \to \Gamma_0$ is taut; and*

(3) *the map $\kappa \circ f\colon W \to \Gamma^*$ is taut.*

Proof. Recall from Theorem 3 that being taut is equivalent to being locally taut. If you replace 'taut' with 'locally taut' in conditions (1)–(3) above, they are easily seen to be equivalent. $\square$

3.8. Continuity. Let $[f]\colon (\Gamma_1, w) \to \Gamma_2$ be a homotopy class of maps between marked graphs. We are interested in how $n^w_{[f]}$ varies as w varies. Define $N_{[f]}\colon \mathcal{W}(\Gamma_1) \to \mathcal{W}(\Gamma_2)$ by $N_{[f]}(w) := n^w_{[f]}$. (There is usually no single $g \in [f]$ that is taut for all w.)

PROPOSITION 3.44. *Let $[f]\colon \Gamma_1 \to \Gamma_2$ be a homotopy class of maps between marked graphs. Then $N_{[f]}\colon \mathcal{W}(\Gamma_1) \to \mathcal{W}(\Gamma_2)$ is continuous and PL.*

Proof. Pick $w_0 \in \mathcal{W}(\Gamma_1)$ and $e_0 \in \mathrm{Edge}(\Gamma_2)$, and let $K = n^{w_0}_{[f]}(e_0)$. Suppose that $w \in \mathcal{W}(\Gamma_1)$ is some set of weights with $|w(e) - w_0(e)| < \varepsilon$ for all e. To show continuity of $N_{[f]}$ near w_0, we will give upper and lower bounds on $n^w_{[f]}(e_0)$ in terms of K and ε.

To get an upper bound, let $f_0 \in [f]$ be a taut map from (Γ_1, w_0) to Γ_2. Pick a regular value $y \in e_0$ of f_0, and let $k = |f_0^{-1}(y)|$. Then since $f_0 \in [f]$,

$$n^w_{[f]}(e_0) \leqslant n^w_{f_0}(y) \leqslant n^{w_0}_{f_0}(y) + k\varepsilon = K + k\varepsilon,$$

as desired.

To get a lower bound, let $E_1 = \{e \in \mathrm{Edge}(\Gamma_1) \mid w_0(e) \neq 0\}$ and $v = \min_{e \in E_1} w_0(e)$. Make sure $\varepsilon < v$, and let $M = K/(v - \varepsilon)$. Let $g\colon (\Gamma_1, w) \to \Gamma_2$

be a taut map in $[f]$. Now $n_g^w(e_0)$ can be written as an integer linear combination of weights of edges:

$$n_g^w(e_0) = \sum_{e \in \text{Edge}(\Gamma_1)} a_e w(e).$$

If $n_g^w(e_0) \geqslant K$, we are done. Otherwise, we must have $\sum_{e \in E_1} a_e \leqslant M$, and so

$$n_g^w(e_0) \geqslant \sum_{e \in E_1} a_e w(e)$$

$$\geqslant -M\varepsilon + \sum_{e \in \text{Edge}(\Gamma_1)} a_e w_0(e)$$

$$\geqslant -M\varepsilon + K.$$

The last inequality uses $n_g^{w_0}(e_0) \geqslant K$, which follows from the definition of K.

To show that $N_{[f]}$ is PL, by continuity we may restrict attention to positive weights $w \in \mathcal{W}^+(\Gamma_1)$. Then $n_{[f]}^w(e)$ can be computed by taking a taut representative $g \in [f]$ and finding $g^{-1}(x)$ for x a generic point on e close to one of the endpoints. The combinatorial type of g determines $g^{-1}(x)$, and by Proposition 3.38 there are only finitely many possibilities for $g^{-1}(x)$. For each possibility, $n_g(e)$ is an integer linear combination of weights on Γ_1. By definition $n_{[f]}^w(e)$ is the minimum of these possibilities, so $N_{[f]}$ is a minimum of a finite set of linear functions. $\qquad\square$

4. Weak graphs and minimizing Lipschitz energy

We start by recalling White's result that the minimal Lipschitz stretch factor of a map between length graphs is given by the maximal ratio of lengths of curves.

THEOREM 4 (White). *Let $[\phi]\colon K_1 \to K_2$ be a homotopy class of maps between marked length graphs. Then there is a representative $\psi \in [\phi]$ and a marked curve $c\colon C \to K_1$ so that the sequence*

$$C \xrightarrow{\ c\ } K_1 \xrightarrow{\ \psi\ } K_2$$

is tight. In particular,

$$\text{Lip}[\phi] = \text{Lip}(\psi) = \frac{\ell[\psi \circ c]}{\ell[c]} = \sup_{c\colon C \to K_1} \frac{\ell[\psi \circ c]}{\ell[c]},$$

where the supremum runs over all nontrivial curves. We get the same quantity if we take the supremum over multicurves or over curves that cross each edge of K_1 at most twice.

A version of Theorem 4 appears in papers by Bestvina [**Bes11**, Proposition 2.1] and Francaviglia–Martino [**FM11**, Proposition 3.11], and is attributed to Tad White (unpublished). Both papers work in the context of outer space and so assume that ϕ is a homotopy equivalence, although that assumption is never used in the proof. The extension to marked graphs is also immediate; we omit the proof.

We need a less trivial extension of Theorem 4: we will need to understand the behavior of energies when lengths degenerate to 0. This leads us to *weak graphs*, in which edges may have length 0. The definition of maps between weak graphs requires a little care, since we need to remember homotopy information that may not be present at the level of the graph obtained by collapsing all 0-length edges.

DEFINITION 4.1. A *weak graph* is a graph Γ in which certain edges are designated as *null edges*; the union of the null edges forms the *null graph*. The *collapsed graph* Γ^* of Γ is the graph obtained by identifying each null edge to a single point. There is a natural *collapsing map* $\kappa \colon \Gamma \to \Gamma^*$. If Γ is a marked graph, then Γ^* is also marked, with marked points the image of the marked points of Γ. Observe that κ is a homotopy equivalence if and only if the null graph is a forest, with at most one marked point in each component of the null graph.

If (Γ, ℓ) is a weak length graph, we consider Γ to be a weak graph, where the null edges are the edges of length 0. The lengths $\ell(e)$ determine a pseudometric on Γ and a metric on Γ^*.

If Γ_1 and Γ_2 are weak graphs, a PL *weak map* from Γ_1 to Γ_2 is a pair $(\widetilde{f}, f)$ of a PL map $f \colon \Gamma_1^* \to \Gamma_2^*$ between the collapsed graphs, together with a map $\widetilde{f} \colon \Gamma_1 \to \Gamma_2$ that is a local homotopy lift of f, in the following sense. Pick disjoint regular neighborhoods N_v^* of each vertex v of Γ_2^*. Let $N_2^* = \bigcup_v N_v^* \subset \Gamma_2^*$, $N_1^* = f^{-1}(N_2^*)$, $N_2 = \kappa_2^{-1}(N_2^*)$, and $N_1 = \kappa_1^{-1}(N_1^*)$. Then we require that $N_1 = \widetilde{f}^{-1}(N_2)$ and that the diagram

$$\begin{array}{ccc} \Gamma_1 & \xrightarrow{\ \widetilde{f}\ } & \Gamma_2 \\ \big\downarrow{\scriptstyle \kappa_1} & & \big\downarrow{\scriptstyle \kappa_2} \\ \Gamma_1^* & \xrightarrow{\ f\ } & \Gamma_2^* \end{array} \qquad (4.2)$$

commutes up to a homotopy that is the identity on $\Gamma_1 \backslash N_1$. We also call f the *shadow* of $\widetilde{f}$.

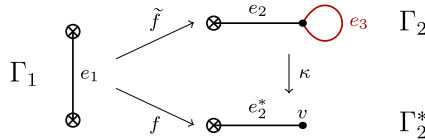If the Γ_i are marked, we also require that f and $\widetilde{f}$ be marked maps.

Figure 4. A weak map where an exact lift is impossible. On the right is a weak graph Γ_2 with a null edge e_3, and its collapsed graph Γ_2^*. The map $f : \Gamma_1 \to \Gamma_2^*$ maps e_1 forward and backward e_2^*, so that the inverse image of the vertex v is a single point. The local lift $\widetilde{f} : \Gamma_1 \to \Gamma_2$ maps e_1 along e_2, around e_3, and then back along e_2. There is no exact lift of f.

We do not require that $\widetilde{f}$ be an exact lift of f, as those do not always exist; see Figure 4.

By the *energy* of a weak map $(\widetilde{f}, f)$ (with additional weight, elastic, or length structure on the graphs, as appropriate), we mean the energy of the map f between collapsed graphs, using the same analytic expression as for maps between ordinary graphs. This applies to all of the energies in Definition 1.30.

PROPOSITION 4.3. *Let Γ_1, Γ_2 be weak graphs, and let $[g] : \Gamma_1 \to \Gamma_2$ be a homotopy class of maps between the uncollapsed graphs. Then there is a weak map $(\widetilde{f}, f)$ between Γ_1 and Γ_2, with $\widetilde{f} \in [g]$, if and only if:*

- *there is no null cycle in Γ_1 that maps to a nonnull cycle in Γ_2; and*

- *there is no null path in Γ_1 between marked points that maps to a nonnull path in Γ_2.*

Here a *null cycle* or *null path* means one in which every edge is null.

Proof. The 'only if' direction is immediate. The other direction is an application of Theorem 4, applied to a metric on the Γ_i in which all nonnull edges have length 1 and all null edges have sufficiently small length. □

In light of Proposition 4.3, we make the following definition.

DEFINITION 4.4. Let Γ_1, Γ_2 be marked weak graphs. A *homotopy class* of maps between them is a homotopy class of maps between the uncollapsed graphs that contains at least one weak map $(f, \widetilde{f})$, that is, so that (up to homotopy) the image of every null cycle (respectively path) is a null cycle (respectively path). Note that this is the homotopy class of $\widetilde{f}$; the homotopy class of f typically has less information.

We next generalize the notions of reduced maps.

DEFINITION 4.5. Let $(\widetilde{f}, f)\colon (\Gamma_1, \Gamma_1^*) \to (\Gamma_2, \Gamma_2^*)$ be a weak marked map. Suppose we have point $y \in \Gamma_2^*$, a regular neighborhood N_y^* of y, and a component Z of $\kappa_1^{-1}(f^{-1}(N_y^*))$ that is not an entire component of Γ_1. Then we say that Z is a *dead end* if $\widetilde{f}$ can be changed by a local homotopy so that $\widetilde{f}(Z)$ does not intersect $\kappa_2^{-1}(y)$. (Compare Definition 3.3.) Concretely, we have the following, where $N_y = \kappa_2^{-1}(N_y^*)$.

(1) If Z contains a marked point, then Z is not a dead end.

(2) If there are two points $x_1, x_2 \in \partial Z$ so that $\widetilde{f}(x_1)$ and $\widetilde{f}(x_2)$ are distinct points of ∂N_y, then Z is not a dead end.

(3) If there are points $x_1, x_2 \in \partial Z$ (possibly identical) and a path γ in Z between them so that $\widetilde{f}(x_1) = \widetilde{f}(x_2)$ and $\widetilde{f}(\gamma)$ is nontrivial in $\pi_1(N_y, \widetilde{f}(x_1))$, then Z is not a dead end.

(4) Otherwise, Z is a dead end.

If $(\widetilde{f}, f)$ has no dead ends, we say that it is *reduced*. As in Definition 3.3, we can *reduce* dead ends: if Z is a dead end mapping to $y \in \Gamma_2^*$, modify f so that $f(\kappa_2(Z))$ misses y and continue to pull until you reach another singular value of f.

PROPOSITION 4.6. *Every weak map is homotopic to a reduced map. Reduction does not increase E_q^p for any $p \leqslant q$.*

Proof. As in Proposition 3.5, energies are not increased as you reduce dead ends. To find the reduced map, as in Proposition 3.4, repeatedly reduce dead ends. $\square$

Furthermore, for a weak curve (a weak map from a marked 1-manifold to a weak graph), reduced maps minimize all energies E_p^1, as in Lemma 3.10 and Proposition 3.8.

THEOREM 4′. *Let $[\phi]\colon K_1 \to K_2$ be a homotopy class of maps between weak marked length graphs, so that K_1 has at least one nonnull edge. Then there is a weak map $\psi \in [\phi]$ and a marked curve $c\colon C \to K_1$ with $\ell[c] > 0$ so that the sequence*

$$C \xrightarrow{\;c\;} K_1 \xrightarrow{\;\psi\;} K_2$$

is tight. In particular,

$$\mathrm{Lip}[\phi] = \mathrm{Lip}(\psi) = \frac{\ell[\psi \circ c]}{\ell[c]} = \sup_{\substack{c' : C \to K_1 \\ \ell[c] > 0}} \frac{\ell[\psi \circ c]}{\ell[c]},$$

where the supremum runs over all curves c' of positive length in K_1. We get the same quantity if we take the supremum over multicurves or over curves that cross each edge of K_1 at most twice.

To prove Theorem 4′, we restrict the maps we consider, first for ordinary (nonweak) maps.

DEFINITION 4.7. A PL map $f : K_1 \to K_2$ between marked length graphs is *constant-derivative* if it is edge-reduced and $|f'|$ is constant on each edge of K_1. A constant-derivative map is determined by its combinatorial type (Definition 3.37) plus a bounded amount of continuous data: for each vertex v of K_1 that maps to the interior of an edge e of K_2, record where $f(v)$ is along e.

LEMMA 4.8. *For any $D > 0$, there are only finitely many combinatorial types of constant-derivative maps among maps with Lipschitz constant $\leqslant D$.*

Proof. The bound on derivatives gives a bound on how many edges of K_2 the image of an edge of K_1 can cross. □

Since Lipschitz energy minimizers are constant-derivative, the above lemma reduces us to finitely many cases to consider. We can give parallel definitions in the case of weak maps.

DEFINITION 4.9. A weak map $(\widetilde{f}, f) : \Gamma_1 \to \Gamma_2$ is *edge-reduced* if $\widetilde{f}$ is edge-reduced in the sense of Definition 3.1. The *combinatorial type* of an edge-reduced weak map is the combinatorial type of f. (The map f need not itself be edge-reduced; see Figure 4 for an example. There is still a well-defined sequence of oriented edges that the image passes over, so the extension of Definition 3.37 is clear.) If in addition Γ_1 and Γ_2 have length structures, we say that $(\widetilde{f}, f)$ is *constant-derivative* if f is constant-derivative.

REMARK 4.10. Definition 4.9 depends only on the shadow f of the weak map. As a result, unlike for ordinary maps, the combinatorial type of a weak map $(\widetilde{f}, f)$ does not determine its homotopy class $[\widetilde{f}]$. Furthermore, for a constant-derivative weak map $(\widetilde{f}, f)$, if we fix the combinatorial type of f, the homotopy

class $\big[\widetilde{f}\big]$, and the continuous data of the position $f(v)$ for $v \in \mathrm{Vert}(\varGamma_1)$, we still may not determine $(\widetilde{f}, f)$ up to local homotopy. Nevertheless, since the energies we consider only depend on f, we are able to work with this definition.

Proof of Theorem $4'$. We follow the previous proofs [**Bes11**, **FM11**], adapted to weak maps.

First, we show there is a map that minimizes Lipschitz energy. The condition that no null cycle of K_1 maps to a nonnull cycle of K_2 guarantees that there is some constant-derivative weak map $(\widetilde{\psi}_0, \psi_0)$ of finite Lipschitz energy. Then any optimizer $(\widetilde{\psi}, \psi)$ must have Lipschitz energy less that $\mathrm{Lip}(\psi_0)$. By Lemma 4.8 there are only finitely many combinatorial types of weak maps $(\widetilde{\psi}, \psi)$ to consider. For each constant-derivative maps of a fixed combinatorial type, $\mathrm{Lip}(\psi)$ is a continuous function of the position of the endpoints along the target edges, and so achieves its minimum among maps with that type.

Let $(\widetilde{\psi}, \psi)$ be a constant-derivative map realizing this minimum. We next prove the existence of a marked curve c exhibiting the Lipschitz stretch of ψ. For each nonnull edge e of K_1, say that it is:

- a *tension edge* if $|\psi'(e)|$ is the maximal value, $\mathrm{Lip}(\psi)$; and

- a *slack edge* if $|\psi'(e)| < \mathrm{Lip}(\psi)$.

Assume that the set of tension edges of ψ is minimal among maps with minimal $\mathrm{Lip}(\psi)$.

To find the desired curve, we will find a suitable sequence of oriented edges $(\vec{e}_i)_{i=1}^N$ of K_1 passing only over tension and null edges.

LEMMA 4.11. *Let* $(\widetilde{\psi}, \psi)$ *be a constant-derivative map realizing the minimal Lipschitz factor with minimal set of tension edges. Let* $\vec{e}_0$ *be an oriented tension edge of* K_1. *Then we can find a reduced edge-path* $\vec{e}_1, \ldots, \vec{e}_k$ *on* K_1 *so that either:*

- *the* $\vec{e}_i$ *for* $i = 1, \ldots, k$ *are null edges and* $\vec{e}_k$ *ends at a marked point; or*

- *the* $\vec{e}_i$ *for* $i = 1, \ldots, k-1$ *are null edges and* $\vec{e}_k$ *is a tension edge.*

In the first case, we allow for $k = 0$, *if* $\vec{e}_0$ *already ends at a marked point. Furthermore, if* $c_0 \colon [0,1] \to K_1$ *is the map passing over these edges in order, then* $(\widetilde{\psi} \circ c_0, \psi \circ \kappa \circ c_0)$ *is reduced as a weak map, fixing the endpoints.*

Proof. Let v_1 be the end of $\vec{e}_0$. If there is no null edge touching v_1, then either v_1 is marked or there is another tension edge meeting v_1 and mapping by ψ to a different direction at $\psi(v_1)$; in either case we are done. Otherwise, let C_1 be the

connected component of the null graph of K_1 that contains the end of $\vec{e}_0$ and let $\{x\} = \kappa_1(C_1)$. Let $y = \psi(x_1)$ and C_2 be the relevant component of $\kappa_2^{-1}(y)$. C_2 may be a point on an edge of K_2, a vertex of K_2, or a connected null subgraph of K_2. We proceed by cases on C_1 and C_2, parallel to the cases in Definition 4.5.

(1) If C_1 has a marked point x, connect v_1 to x through null edges.

(2) If there is another tension edge $\vec{e}'$ of K_1 incident to C_1 so that $\vec{e}'$ and the reverse of $\vec{e}_0$ map by $\psi \circ \kappa$ to distinct directions in K_2^*, connect $\vec{v}_1$ to the start of $\vec{e}'$ by a reduced edge-path in K_1.

(3) If $\pi_1(C_1)$ maps nontrivially to $\pi_1(C_2)$, connect v_1 to itself by a reduced loop in C_1 that maps to a nontrivial loop in C_2.

If none of the above cases apply, we get a contradiction, as follows. Partition the nonnull edges incident to C_1 according to which direction they map to from y. (Put each edge with $|\psi'| = 0$ into its own part.) Since we are not in cases (1) or (3), we can displace $y \in K_2^*$ in any direction, reducing $|\psi'|$ on one group of edges at the cost of increasing it on all other groups. Since we are not in case (2), we can reduce $|\psi'(\vec{e}_0)|$ so e_0 is no longer a tension edge without creating another tension edge, contradicting the assumption that the set of tension edges of ψ was minimal.

In any of cases (1)–(3), by construction the sequence of edges in K_1 is reduced, and the path in K_2 falls into one of the cases of Definition 4.5 and so is reduced as a weak map. □

To complete the proof of Theorem 4′, start with any tension edge $\vec{e}_0$ of K_1. (There is at least one since not all edges of K_1 are null.) Use Lemma 4.11 to construct a nonbacktracking chain of tension and null edges forward and backward from $\vec{e}_0$. Either the chain ends in a marked point both forward and backward, or there is a repeated oriented edge. In either case we can extract a marked curve c (either an arc or a loop) so that both c and $\psi \circ c$ are reduced (as weak maps). Thus $\mathrm{Lip}(\psi) = \ell(\psi \circ c)/\ell(c) = \ell[\psi \circ c]/\ell[c]$. If we make the null paths as efficient as possible and close off to make a cycle the first time that an oriented edge is repeated, then c will cross each (unoriented) edge of K_1 at most twice. □

REMARK 4.12. The proof of Theorem 4′ is similar to the proof of Proposition 3.19. The presence of null edges introduces extra complications in Theorem 4′.

5. Minimizing Dirichlet energy: harmonic maps

5.1. Definition and existence. We give a concrete description of harmonic maps and prove their existence, making the intuitive description of harmonic maps from the introduction more precise.

DEFINITION 5.1. Let $G = (\Gamma, \alpha)$ be a marked elastic graph and let K be a marked length graph. To any constant-derivative map f, we can associate the *tension-weighted graph* $W_f = (\Gamma, |f'|)$; that is, the weight of an edge e is $|f'(e)|$. Then f is *harmonic* if it is constant-derivative and the associated map $f \colon W_f \to K$ is taut.

For general weighted graphs, it is not immediately obvious when a map is taut. We can simplify the condition in Definition 5.1 to give the triangle inequalities at vertices from the introduction.

PROPOSITION 5.2. *Let $f \colon G \to K$ be a constant-derivative map from a marked elastic graph to a marked length graph. Let $G_0 \subset G$ be the subgraph of G on which $|f'| \neq 0$, and let f_0 be the restriction of f to G_0. Then f is harmonic if and only if $(W_{f_0}, \tau(f_0))$ forms a marked balanced train track in the sense of Definition 3.14.*

Proof. By Lemma 3.40, f is harmonic if and only if $f_0 \colon W_{f_0} \to K$ is taut. If f_0 is taut, the train track in condition (3) of Theorem 3 must be $(W_{f_0}, \tau(f_0))$. $\square$

To be more explicit about the triangle inequalities, fix a constant-derivative map $f \colon G \to K$ from an elastic graph to a length graph. For $v \in G$ and d_K a direction in K from $f(v)$, the *tension* of d_K is

$$T(d_K) = \sum \{|f'(d_G)| \mid d_G \text{ direction in } G \text{ at } v,\ f(d_G) = d_K\}.$$

We say that v is *balanced* if $(T(d_K) \mid d_K \text{ direction in } K \text{ at } f(v))$ satisfies the triangle inequalities.

COROLLARY 5.3. *Let $f \colon G \to K$ be a constant-derivative map from a marked elastic graph to a marked length graph. Then f is harmonic if and only if each unmarked vertex of G is balanced.*

Proof. This is a restatement of Proposition 5.2. $\square$

THEOREM 5. *Let $[f]: G \to K$ be a homotopy class of maps from a marked elastic graph to a marked length graph. Then there is a harmonic map in $[f]$. Furthermore, the following conditions are equivalent.*

(1) *The map f is a global minimum for* Dir.

(2) *The map f is a local minimum for* Dir.

(3) *The map f is harmonic.*

(4) *The natural map $\iota: W_f \to G$ is part of a tight sequence*

$$W_f \xrightarrow{\ \iota\ } G \xrightarrow{\ f\ } K.$$

(5) *There is a weighted multicurve (C, c) on G that forms a tight sequence*

$$C \xrightarrow{\ c\ } G \xrightarrow{\ f\ } K.$$

Proof. We start with the equivalences. By definition, (1) implies (2).

To show that (2) implies (3), let f be a local minimizer for $\mathrm{Dir}(f)$ within $[f]$; we wish to show f is harmonic. One of the first results in calculus of variations is that f is constant-derivative. So by Proposition 5.2 we only need to show the triangle inequalities. Let v be a unmarked vertex of G of valence k, and let $d_1, \ldots, d_k$ be the nonzero directions of K at $f(v)$. For small $\varepsilon > 0$, let $f_{i,\varepsilon}$ be f modified by moving $f(v)$ a distance ε in the direction d_i, extended to the edges so that $f_{i,\varepsilon}$ is still constant-derivative. By hypothesis $\mathrm{Dir}(f) \leqslant \mathrm{Dir}(f_{i,\varepsilon})$. We have

$$\mathrm{Dir}(f_{i,\varepsilon})$$
$$= \mathrm{Dir}(f) + \sum_{d \in f^{-1}(d_i)} -2\varepsilon|f'(d)| + \sum_{j \neq i} \sum_{d \in f^{-1}(d_j)} 2\varepsilon|f'(d)| + \sum_{d \text{ direction at } v} \varepsilon^2/\alpha(d).$$

To be a local minimum, we must have for each i

$$\frac{d}{d\varepsilon}\left(\mathrm{Dir}(f_{i,\varepsilon})\right)\bigg|_{\varepsilon=0} \geqslant 0.$$

This gives the i'th triangle inequality at v, so f is harmonic.

To see that (3) implies (4), suppose that f is harmonic. Then $f \circ \iota$ is taut and therefore energy-minimizing. Furthermore,

$$\mathrm{EL}(\iota) = \sum_{e \in \mathrm{Edge}(\Gamma)} \alpha(e)|f'(e)|^2 = \mathrm{Dir}(f)$$

$$\ell(f \circ \iota) = \sum_{e \in \text{Edge}(\Gamma)} |f'(e)| \, \ell(f(e)) = \text{Dir}(f),$$

so the energies are multiplicative, as desired.

Proposition 3.19 tells us that (4) implies (5).

Lemma 1.34 tells us that (4) or (5) implies (1).

To prove existence, let g be any constant-derivative map. Then $\text{Dir}[f] \leqslant \text{Dir}(g)$. For each edge e, this gives an upper bound on possible values of $|f'(e)|$ for any minimizer. Then, as in the proof of Theorem $4'$, there are only finitely many combinatorial types below this bound, and for each combinatorial type there is a compact space of PL maps. There is therefore a global minimum for $\text{Dir}(f)$, which is harmonic. □

5.2. Harmonic maps to weak graphs.

As with Lipschitz energy, we need a generalization of Theorem 5 to allow the target to be a weak length graph.

DEFINITION 5.4. Let $(\widetilde{f}, f) \colon \Gamma_1 \to \Gamma_2$ be weak map between marked weak graphs, and suppose there is a weight structure w_1 on Γ_1. Then we say that $(\widetilde{f}, f)$ is *taut* if there is a lift $\widetilde{f'}$ locally homotopic to $\widetilde{f}$ so that $\widetilde{f'}$ is taut.

DEFINITION 5.5. Let $G = (\Gamma, \alpha)$ be a marked elastic graph and let K be a marked weak length graph. We say that a weak map $(\widetilde{f}, f) \colon G \to K$ is *harmonic* if it is constant-derivative and the weak map $W_f \to K$ from the tension-weighted graph is taut.

We do not allow G to be a weak graph. The definition could be extended to this case with some more work (since when e is a null edge, $|f'(e)|$ is not defined), but we will not need it.

We do not know a concise version of Corollary 5.3 in this setting.

THEOREM $5'$. *Let $[f] \colon G \to K$ be a homotopy class of maps from a marked elastic graph to a marked weak length graph. Then there is a harmonic weak map in $[f]$. Furthermore, the following conditions are equivalent.*

(1) *The weak map f is a global minimum for* Dir.

(2) *The weak map f is a local minimum for* Dir.

(3) *The weak map f is harmonic.*

(4) *There is a weighted multicurve (C, c) on G that forms a tight sequence*

$$C \xrightarrow{c} G \xrightarrow{f} K.$$

Proof. The most significant change from the proof of Theorem 5 is the proof that (2) implies (3), which we now do. Let $(\widetilde{g}, g)$ be a constant-derivative PL map, and suppose that g is not harmonic in the sense of Definition 5.5. Then we will find a local modification that reduces $\mathrm{Dir}(g)$. The possible local modifications are more complicated than in the proof of Theorem 5; we use the technology of taut maps from Section 3.

Fix ε sufficiently small. For v a vertex of G, let $N_v^* \subset K^*$ be the ε-neighborhood of $f(v)$. Make sure ε is small enough so that the N_v^* are disjoint regular neighborhoods. Let $N^* = \bigcup_v N_v^*$, $N_v = \kappa^{-1}(N_v^*) \subset K$, $N = \bigcup_v N_v$, $M_v = g^{-1}(N_v^*) \subset G$, and $M = \bigcup_v M_v$. We will use M and N for our neighborhoods in the definition of a weak map. On the restriction of $\widetilde{g}$ to each neighborhood, set up a marked local model, analogously to Definition 3.30, except that the target model is not necessarily a star. Using Theorem 3, choose the local lift $\widetilde{g}$ so each local model is taut as a map from W_g.

By assumption, $\widetilde{g}$ is not taut, and thus is not locally taut in a neighborhood of some point y; necessarily $y \in \partial N_v$ for some vertex v. Let U be a small neighborhood of y, and let Z be a component of $\widetilde{g}^{-1}(U)$ on which $\widetilde{g}$ can be homotoped to reduce $n_{\widetilde{g}}$ (as in Definition 3.6 on locally taut). The exterior edges of Z (those that meet $\widetilde{g}^{-1}(U)$) can be divided into *inside* edges, whose images under $\widetilde{g}$ point away toward v, and *outside* edges, whose images point away from v.

Since the marked local models for $\widetilde{g}$ are all taut, the total tension of the inside edges cannot be reduced; in particular, the total tension of the inside edges is less than or equal to the total tension of the outside edges. In fact, since $\widetilde{g}$ as a whole is not locally taut at y, this inequality is strict:

$$\sum_{e \text{ inside}} |g'(e)| < \sum_{e \text{ outside}} |g'(e)|. \tag{5.6}$$

Let $(\widetilde{h}, h)$ be the constant-derivative small modification of $(\widetilde{g}, g)$ so that h maps each vertex in Z to y and agrees with g on all other vertices. In h the length of the image of each inside edge was increased by ε, and the length of the image of each outside edge was decreased by ε. Then

$$\mathrm{Dir}(h) - \mathrm{Dir}(g) = \sum_{e \text{ inside}} \frac{\ell(g(e))\varepsilon + \varepsilon^2}{\alpha(e)} + \sum_{e \text{ outside}} \frac{-\ell(g(e))\varepsilon + \varepsilon^2}{\alpha(e)}$$

$$= \varepsilon \left(\sum_{e \text{ inside}} |f'(e)| - \sum_{e \text{ outside}} |f'(e)| \right) + O(\varepsilon^2),$$

which is negative for ε sufficiently small by Equation (5.6). Thus $\mathrm{Dir}(h) < \mathrm{Dir}(g)$ and g is not a local minimum for Dir.

The rest of the proof is almost unchanged: (1) implies (2) by definition, and (4) implies (1) by properties of tight sequences. For (3) implies (4), if f is harmonic, by Theorem 3 there is a weighted multicurve (C, c) on G so that $n_c = |f'|$. Then $C \xrightarrow{\;c\;} G \xrightarrow{\;f\;} K$ is tight. Finally, existence follows from by bounding the combinatorial types as before. □

We can improve the local lifts in a weak harmonic map a little.

DEFINITION 5.7. In a weak map $(\widetilde{f}, f)\colon \Gamma_1 \to \Gamma_2$, we say that the local lift $\widetilde{f}$ is *vertex-precise* if, for every vertex v of Γ_1, $\kappa_2(\widetilde{f}(v)) = f(\kappa_1(v))$ on the nose (with no need for homotopy).

PROPOSITION 5.8. *If $(\widetilde{f}, f)\colon G \to K$ is a harmonic weak map, then $\widetilde{f}$ can be chosen to be vertex-precise and taut as a map from W_f to K.*

Proof. By definition of harmonic weak maps, we can find a taut initial lift $\widetilde{f}$. If $\widetilde{f}$ does not map vertices to vertices, pick some vertex v so that $f(v) \neq \kappa(\widetilde{f}(v))$. Then $\widetilde{f}(v)$ lies on an edge e incident to $\kappa^{-1}(f(v))$. Since $\widetilde{f}$ is taut, $n_{\widetilde{f}}$ is constant on e. We can thus push $\widetilde{f}(v)$ into $\kappa^{-1}(f(v))$ without changing f or $n_{\widetilde{f}}$ (and thus keeping $\widetilde{f}$ taut). Repeat until $\widetilde{f}$ maps vertices to vertices. □

5.3. Uniqueness and continuity.

Harmonic maps are not in general unique in their homotopy class. For instance, if the target length graph is a circle, then composing a harmonic map with any rotation of the circle gives another harmonic map. However, the length of the image of an edge in a harmonic map is unique. In fact, the lengths are unique in a larger set.

DEFINITION 5.9. For Γ a marked graph, K a weak marked length graph, and $[f]\colon \Gamma \to K$ a homotopy class of maps, a *relaxed map r* with respect to $[f]$ is an assignment of a length $r(e)$ to each edge e of Γ, so that, for any taut weighted marked multicurve (C, c) on Γ,

$$\sum_{e \in \mathrm{Edge}(\Gamma)} n_c(e) r(e) \geqslant \ell[f \circ c]. \tag{5.10}$$

A relaxed map r naturally gives a weak length metric on Γ. Let $\mathrm{Rlx}[f] \subset \mathcal{L}(\Gamma)$ be the set of relaxed maps. We write $\mathrm{Rlx}_\ell[f]$ if we need to make precise the dependence on $\ell \in \mathcal{L}(K)$.

Although a relaxed map is not, in fact, any sort of map, the next three lemmas show how relaxed maps are related to actual maps.

LEMMA 5.11. *If $[f]: \Gamma_1 \to \Gamma_2$ is a homotopy class, $r \in \mathcal{L}(\Gamma_1)$, and $\ell \in \mathcal{L}(\Gamma_2)$, then $r \in \mathrm{Rlx}_\ell[f]$ if and only if there is a map $g \in [f]$ with $\mathrm{Lip}_\ell^r(g) \leqslant 1$.*

Proof. This is Theorem 4′. □

LEMMA 5.12. *If $f: \Gamma \to K$ is a weak PL map, then $f^*\ell \in \mathrm{Rlx}_\ell[f]$.*

Proof. By definition, $\mathrm{Lip}_\ell^{f^*\ell}(f) = 1$. Apply Lemma 5.11. □

LEMMA 5.13. *If $[f]: \Gamma \to K$ is a homotopy class of maps from a marked graph Γ to a nontrivial weak marked length graph K and $r \in \mathrm{Rlx}[f]$, then there is a PL map $g \in [f]$ so that $r = g^*\ell$.*

Proof. Lemma 5.11 gives a PL map $g_0: (\Gamma, r) \to K$ in $[f]$ with $\mathrm{Lip}(g_0) \leqslant 1$. That is, $\ell(g_0(e)) \leqslant r(e)$ for each edge e of Γ. Define g by modifying g_0: for each edge e on which $\ell(g_0(e)) < r(e)$, make the length of the image of e longer by introducing some zigzag folds. □

LEMMA 5.14. *Definition 5.9 does not change if we let (C, c) be:*

(1) *a marked balanced train track;*

(2) *a marked weighted multicurve (as in Definition 5.9);*

(3) *a marked curve with weight 1; or*

(4) *a marked curve with weight 1 that crosses each edge at most twice.*

Proof. The types of curve-like objects are progressively more restrictive, so we need to show that the existence of a structure of one type violating Equation (5.10) implies the next. Condition (1) implies condition (2) by Proposition 3.19. Condition (2) implies condition (3) by the additivity of Equation (5.10) over connected components. Condition (3) implies condition (4) by Lemma 5.11 and Theorem 4′ or, more simply, by taking any curve, looking for a maximal segment with no repeated oriented edges, doing cut and paste if necessary, and then using additivity over the connected components. □

LEMMA 5.15. *For $[f]\colon \Gamma_1 \to \Gamma_2$ a homotopy class of maps between marked graphs and $\ell \in \mathcal{L}(\Gamma_2)$, $\mathrm{Rlx}_\ell[f]$ is a closed, noncompact, convex polytope defined by finitely many inequalities, each inequality depending linearly on ℓ.*

Proof. This follows from condition (4) of Lemma 5.14, as there are only finitely many curves crossing each edge at most twice, and Equation (5.10) cuts out a closed half-space for each such curve. Scaling a relaxed map by a factor $\lambda > 1$ gives another relaxed map, so $\mathrm{Rlx}[f]$ is not compact. □

If $r \in \mathrm{Rlx}[f]$ is a relaxed map where the domain is an elastic graph (Γ, α), we can define the Dirichlet energy of r:

$$\mathrm{Dir}^\alpha(r) := \sum_{e \in \mathrm{Edge}(\Gamma)} \frac{r(e)^2}{\alpha(e)}. \tag{5.16}$$

In fact, Equation (5.16) makes sense for any $r \in \mathcal{L}(\Gamma)$.

We can now give the uniqueness statement.

PROPOSITION 5.17. *If $f\colon G \to K$ is a harmonic map from a marked elastic graph to a weak marked length graph, then $f^*\ell$ uniquely minimizes Dirichlet energy on $\mathrm{Rlx}_\ell[f]$.*

Proof. Let $G = (\Gamma, \alpha)$. The function Dir^α is strictly convex on $\mathcal{L}(\Gamma)$, and its sublevel sets are compact. As such, Dir^α achieves a unique minimum on the closed convex set $\mathrm{Rlx}[f]$. Since f was harmonic, it minimizes $\mathrm{Dir}(f)$ in $[f]$; since every point in $\mathrm{Rlx}[f]$ gives the lengths of an actual map in $[f]$, the minimizer in $\mathrm{Rlx}[f]$ must be $f^*\ell$. □

In light of Proposition 5.17, we can think of Dirichlet energy and the length of edges of the harmonic minimizer as functions on $\mathcal{L}(K)$.

DEFINITION 5.18. For $[f]$ a homotopy class of maps from a marked elastic graph (Γ_1, α) to a marked graph Γ_2, define

$$\mathrm{Dir}_{[f]}\colon \mathcal{L}(\Gamma_2) \to \mathbb{R}$$
$$H_{[f]}\colon \mathcal{L}(\Gamma_2) \to \mathcal{L}(\Gamma_1)$$

by setting $\mathrm{Dir}_{[f]}(\ell)$ to $\mathrm{Dir}^\alpha_\ell[f]$ and $H_{[f]}(\ell)$ to the relaxed map in $\mathrm{Rlx}_\ell[f]$ minimizing Dir^α.

PROPOSITION 5.19. *Let* $[f]\colon G \to \Gamma_2$ *be a homotopy class of maps from a marked elastic graph to a marked graph. Then* $\mathrm{Dir}_{[f]}$ *and* $H_{[f]}$ *are continuous functions on* $\mathcal{L}(\Gamma_2)$, *with* $\mathrm{Dir}_{[f]}$ *piecewise-quadratic and* $H_{[f]}$ *piecewise-linear.*

Proof. As $\ell \in \mathcal{L}(\Gamma_2)$ varies, the closed convex set $\mathrm{Rlx}_\ell[f]$ varies continuously by Lemma 5.15. Since Dir^α is strictly convex, both the value and location of the minimum of Dir^α on $\mathrm{Rlx}_\ell[f]$ depend continuously on ℓ.

Furthermore, since Dir^α is quadratic on $\mathcal{L}(\Gamma)$ and $\mathrm{Rlx}_\ell[f]$ depends piecewise-linearly on ℓ, the value of the minimum of Dir^α on $\mathrm{Rlx}_\ell[f]$ is a piecewise-quadratic function of ℓ and the location of the minimum is a PL function of ℓ. (The particular quadratic or linear function depends on the face of $\mathrm{Rlx}_\ell[f]$ containing the minimum.) □

See also Remark 6.33.

REMARK 5.20. An alternate way to see that $\mathrm{Dir}_{[f]}$ is piecewise-quadratic and that $H_{[f]}$ is PL is to note that they are respectively quadratic and linear for a fixed combinatorial type of a harmonic representative, and only finitely many combinatorial types of maps appear by Theorem 5′ and Proposition 3.38. The combinatorial type is related to the face of $\mathrm{Rlx}_\ell[f]$ containing $H_{[f]}(\ell)$.

6. Minimizing embedding energy

6.1. Characterizing minimizers: λ-filling maps. We now turn to the proof of Theorem 1, starting with a characterization of which maps can appear as minimizers of $\mathrm{Emb}[\phi]$. Since we will show that a map that minimizes $\mathrm{Emb}[\phi]$ is a simultaneous minimizer of the ratio of Dirichlet energy (for maps to a length graph) and of the ratio of extremal length (from maps from a weighted multicurve), it is helpful to have a notion that encapsulates both weights and lengths to prove both tightness statements uniformly.

DEFINITION 6.1. A *strip graph* $S = (\Gamma, w, \ell)$ is a marked graph Γ with weights $w \in \mathcal{W}(\Gamma)$ and lengths $\ell \in \mathcal{L}(\Gamma)$, so that $w(e) \neq 0$ if and only if $\ell(e) \neq 0$. The strip graph is *positive* if all lengths and weights are positive, and is *balanced* if w is balanced (Definition 3.14).

There is an associated marked weighted graph $W_S = (\Gamma, w)$ and weak marked length graph $K_S = (\Gamma, \ell)$. In addition to these, there is also an associated elastic structure: We say that S is *compatible* with an elastic graph $G_S = (\Gamma, \alpha)$ if $\ell(e) = w(e)\alpha(e)$ for each edge e. (If S is positive, then G_S is unique.) A strip graph also

has an *area*

$$\text{Area}(S) := \sum_{e \in \text{Edge}(\Gamma)} \ell(e) w(e).$$

To build intuition for strip graphs, we relate strip graphs to tight sequences of maps.

PROPOSITION 6.2. *A balanced strip graph S gives a tight sequence of maps*

$$W_S \longrightarrow G_S \longrightarrow K_S \longrightarrow K_S^*.$$

(Recall from Definition 4.1 that K_S^* is K_S with the null edges collapsed.)

Proof. The fact that the map $W_S \to K_S$ is taut (hence energy-minimizing) is the condition that w is balanced. The map $W_s \to K_S^*$ is then taut by Lemmas 3.40 and 3.43. We also have

$$\ell(W_S \to K_S^*) = \text{EL}(W_S \to G_S) = \text{Dir}(G_S \to K_S) = \text{Area}(S)$$
$$\text{Lip}(K_S \to K_S^*) = 1,$$

so the energies are multiplicative, as desired. □

PROPOSITION 6.3. *Let $f : G \to K$ be a harmonic map, with $G = (\Gamma, \alpha)$. Then there is a unique balanced strip structure $S = (\Gamma, w, \ell)$ on Γ, compatible with α, so that f is an isometry from (Γ, ℓ) to K.*

Similarly, let $f : W \to G$ be a taut map from a weighted graph to an elastic graph. Then there is a unique balanced strip structure $S = (\Gamma, w, \ell)$ compatible with G so that $n_f(y) = w(y)$.

Proof. For the first statement, set $\ell(e)$ to be the length of the image, $f(e)$, and set $w(e)$ to be the tension, $|f'(e)|$. For the second statement, set $w(e)$ to be the multiplicity covered $n_f(e)$ and $\ell(e) = \alpha(e) \cdot w(e)$. Verifying the required properties is straightforward. In particular, 'balanced' follows from Corollary 5.3 in the harmonic map case. □

DEFINITION 6.4. Let $S_1 = (\Gamma_1, w_1, \ell_1)$ and $S_2 = (\Gamma_2, w_2, \ell_2)$ be two marked balanced strip graphs. Write G_1 for the elastic graph G_{S_1} associated to S_1, and so forth. (This involves choices if the S_i are not positive.) Let $\phi : \Gamma_1 \to \Gamma_2$ be a PL map. Since there are many different structures on the Γ_i, we give names to ϕ considered as a map between different structures: we have

$$\phi_W^W \colon W_1 \to W_2 \qquad \phi_G^W \colon W_1 \to G_2 \qquad \phi_K^W \colon W_1 \to K_2$$
$$\phi_G^G \colon G_1 \to G_2 \qquad \phi_K^G \colon G_1 \to K_2$$
$$\phi_K^K \colon K_1 \to K_2.$$

Then if $\lambda > 0$ is a real number and S_2 is positive, ϕ is λ-*filling* if:

(1) ϕ_K^W is taut;

(2) lengths are preserved: ϕ_K^K is a local isometry; and

(3) weights are scaled by a factor of λ: for every regular value $y \in \Gamma_2$,

$$\sum_{x \in \phi^{-1}(y)} w_1(x) = \lambda w_2(y). \qquad (6.5)$$

In other words, $n_\phi^{w_1} = \lambda w_2$ and in particular $\mathrm{WR}(\phi_W^W) = \lambda$.

We say that a map ϕ between elastic graphs is λ-filling if there are compatible strip structures so that ϕ is λ-filling as a map between the strip structures.

LEMMA 6.6. *Suppose* $\phi \colon S_1 \to S_2$ *is a* λ-*filling map. Then* $\mathrm{Fill}_{\phi_G^G}$ *is identically equal to* λ. *In particular,* $\mathrm{Emb}(\phi_G^G) = \lambda$.

Proof. Since ϕ is constant on the 0-length edges of S_1, we may delete these edges and assume that S_1 is positive as well. Since ϕ_K^K is length-preserving and $\alpha_i(e) = \ell_i(e)/w_i(e)$ for $i = 1, 2$, it follows that, for regular points $x \in G_1$,

$$\left| \left(\phi_G^G \right)'(x) \right| = w_1(x)/w_2(\phi(x)).$$

The result follows from Equation (6.5). □

LEMMA 6.7. *A* λ-*filling map* $\phi \colon S_1 \to S_2$ *gives a tight sequence of maps*

$$W_1 \xrightarrow{\mathrm{id}_{\Gamma_1}} G_1 \xrightarrow{\phi_G^G} G_2 \xrightarrow{\mathrm{id}_{\Gamma_2}} K_2.$$

Proof. The composite $\phi_K^W \colon W_1 \to K_2$ is taut by assumption. The energies of the various maps are

$\mathrm{EL}(\mathrm{id}_{\Gamma_1}) = \mathrm{Area}(S_1)$ by Proposition 6.2

$\mathrm{Emb}(\phi_G^G) = \lambda$ by Lemma 6.6

$\mathrm{Dir}(\mathrm{id}_{\Gamma_2}) = \mathrm{Area}(S_2) = \mathrm{Area}(S_1)/\lambda$ by Proposition 6.2 and Equation (6.5).

Then $\ell(W_1 \to K_2) = \mathrm{Area}(S_1) = \sqrt{\mathrm{EL}(\mathrm{id}_{\Gamma_1}) \, \mathrm{Emb}(\phi_G^G) \, \mathrm{Dir}(\mathrm{id}_{\Gamma_2})}$, which is multiplicative, as desired. □

PROPOSITION 6.8. *Let $\phi\colon G_1 \to G_2$ be a map between marked elastic graphs G_1 and G_2. If there is a λ-filling map $\psi \in [\phi]$, then $\mathrm{Emb}[\phi] = \mathrm{SF}_{\mathrm{Dir}}[\phi] = \mathrm{SF}_{\mathrm{EL}}[\phi] = \mathrm{Emb}(\psi) = \lambda$.*

Proof. Immediate from Lemmas 6.6, 6.7 and 1.34. □

Thus, λ-filling maps are optimizers for Emb. If there were always a λ-filling map in $[\phi]$, then we would be done with Theorem 1. Unfortunately this is not true.

EXAMPLE 6.9. Let G_1 and G_2 both be the join of two circles, with elastic constants $(2, 2)$ and $(1, 4)$, respectively, and let $\phi\colon G_1 \to G_2$ be the constant-derivative map which is the identity on π_1 and maps the vertex to the vertex:

Then $\mathrm{Emb}(\phi) = 2$. Furthermore, $\mathrm{SF}_{\mathrm{EL}}[\phi] \geqslant 2$, by considering the right-hand loop of G_1. Thus $\mathrm{Emb}[\phi] = \mathrm{Emb}(\phi) = \mathrm{SF}_{\mathrm{EL}}[\phi] = 2$. There is no λ-filling map in $[\phi]$.

DEFINITION 6.10. Let $S_1 = (\Gamma_1, w_1, \ell_1)$ and $S_2 = (\Gamma_2, w_2, \ell_2)$ be two marked balanced strip graphs, with S_2 positive. Then, for $\lambda > 0$, a map $\phi\colon S_1 \to S_2$ is *partially λ-filling* if there are complementary subgraphs $\Delta_i, \Sigma_i \subset \Gamma_i$ (that is, with complementary sets of edges), so that:

(1) $\phi(\Delta_1) \subset \Delta_2$ (so we can write $\phi|_\Delta$ for the restriction);

(2) $\phi(\Sigma_1) \subset \Sigma_2$ (so we can write $\phi|_\Sigma$ for the restriction);

(3) the strip graph structures induced on Δ_i and Δ_2 are balanced;

(4) the restriction $\phi|_\Delta\colon \Delta_1 \to \Delta_2$ is λ-filling with respect to the induced strip graph structures;

(5) ϕ is everywhere length-preserving; and

(6) $\phi|_\Sigma$ scales weights by strictly less than λ: for every regular value $y \in \Sigma_2$,

$$\sum_{x \in \phi^{-1}(y)} w_1(x) < \lambda w_2(y).$$

We call Δ_1 and Δ_2 the *filling subgraphs* of S_1 and S_2. Set W_1^0 to be the marked balanced weighted graph (Δ_1, w_1) and set K_2^{0*} to be the marked length graph (Γ_2, ℓ_2) with the edges in Σ_2 collapsed to points.

LEMMA 6.11. *If $\phi\colon S_1 \to S_2$ is a partially λ-filling map then, with notation as in Definition 6.10, there is a tight sequence*

$$W_1^0 \longrightarrow G_1 \xrightarrow{\phi_G^G} G_2 \longrightarrow K_2^{0*}.$$

Here, we are thinking about K_2^{0*} as an ordinary graph. (Later there will be an associated weak length graph.)

Proof. The version of Lemma 6.6 in this context says that $\mathrm{Fill}_{\phi_G^G}(y) = \lambda$ if y is a regular point in Δ_2 and $\mathrm{Fill}_{\phi_G^G}(y) < \lambda$ if y is a regular point in Σ_2. We then have

$$\mathrm{EL}(W_1^0 \to G_1) = \mathrm{Area}(\Delta_1)$$
$$\mathrm{Emb}(G_1 \to G_2) = \lambda$$
$$\mathrm{Dir}(G_2 \to K_2^{0*}) = \mathrm{Area}(\Delta_2) = \mathrm{Area}(\Delta_1)/\lambda$$
$$\ell\big[W_1^0 \to K_2^{0*}\big] = \ell(W_1^0 \to K_2^{0*}) = \mathrm{Area}(\Delta_1),$$

which is multiplicative.

Since $\phi|_\Delta$ is λ-filling, the map $W_1^0 \to \Delta_2$ is taut. By Lemma 3.43 applied to $W_1^0 \to \Gamma_2$, this implies that $W_1^0 \to K_2^{0*}$ is taut, completing the proof the sequence is tight. $\square$

Thus partially λ-filling maps are also global minimizers for Emb. The bulk of this section will be devoted to proving they exist:

PROPOSITION 6.12. *For $[\phi]\colon G_1 \to G_2$ a homotopy class of maps between marked elastic graphs, there are compatible strip structures on G_1 and G_2 and map $\psi \in [\phi]$ so that ψ is partially λ-filling.*

Proposition 6.12 suffices to prove Theorem 1. We spell out some further consequences.

PROPOSITION 6.13. *For $[\phi]\colon G_1 \to G_2$ a homotopy class of maps between elastic graphs, there is a tight sequence*

$$C \xrightarrow{c} T \xrightarrow{t} G_1 \xrightarrow{\psi} G_2 \xrightarrow{f} K$$

where $\psi \in [\phi]$, T is a balanced train track whose underlying graph is a subgraph of G_1, t is the inclusion map, K is a length graph whose underlying graph is obtained by collapsing some edges of G_2, f is the collapsing map, (C, c) is a weighted multicurve saturating T, and $\psi \circ t$ is a train-track map. Furthermore, the image of $\psi \circ t$ is the union of the edges that are not collapsed by f, and Fill_ψ is constant and maximal on those edges.

Proof of Proposition 6.13, *assuming Proposition* 6.12. By Proposition 6.12, there is a partially λ-filling map $\psi \in [\phi]$. Lemma 6.11 gives a tight sequence

$$W_1^0 \xrightarrow{\ i\ } G_1 \xrightarrow{\ \psi\ } G_2 \xrightarrow{\ f\ } K_2^{0*}.$$

Since i is taut, Theorem 3 gives a tight sequence of maps

$$C \xrightarrow{\ c\ } T \xrightarrow{\ t\ } W_1^0 \xrightarrow{\ i\ } G_1 \xrightarrow{\ \psi\ } G_2 \xrightarrow{\ f\ } K_2^{0*}$$

where C is a weighted multicurve and T is a balanced train track, whose underlying graph is a subgraph of W_1^0 (and so G_1). Dropping W_1^0 gives the desired sequence. The edges not collapsed by f, the image of $\psi \circ i \circ t$, and the edges where Fill_ψ is constant are all equal to the filling subgraph of G_2. $\qquad \square$

Proof of Theorem 1, *assuming Proposition* 6.12. Immediate consequence of Proposition 6.13 and Lemma 1.34. $\qquad \square$

6.2. Iterating to optimize embedding energy.

One approach to proving Proposition 6.12 would be to study those maps that locally minimize the embedding energy, analogously to the proof of Theorem 4. From Lagrange multipliers at the local minimum, you can extract lengths and weights to form the desired strip structure. This is a little involved, since the relevant ambient space in which to apply the theory of Lagrange multipliers is not obvious; there is no obvious analogue of the space of constant-derivative maps used in Section 5. We will therefore take a different approach, one that also suggests an algorithm to actually compute the embedding energy.

From a homotopy class $[\phi] \colon G_1 \to G_2$ of maps between marked elastic graphs, we will give an iteration that has a fixed point at a partially λ-filling map. To motivate the iteration, we give an analogue in the setting of vector spaces. First, recall that a strictly convex norm on a finite-dimensional vector space V defines a *dual norm* on the dual space V^*. This is the minimal norm that satisfies, for all $v \in V$ and $v^* \in V^*$,

$$\langle v^*, v \rangle \leqslant \|v\| \|v^*\|. \tag{6.14}$$

Equation (6.14) is tight in the sense of Proposition 2.15, namely, for every nonzero $v \in V$ there is a nonzero $v^* \in V^*$, unique up to scale, so that Equation (6.14) is an equality. If in addition $\|v^*\| = \|v\|$, we say that v and v^* *support* each other. (The hyperplane corresponding to v^* is parallel to a supporting hyperplane for a norm ball in V at v.)

EXAMPLE 6.15. If $\|v\| = \sqrt{\langle v, v \rangle}$ for an inner product $\langle \cdot, \cdot \rangle$, the map from a vector to its supporting vector is the canonical isomorphism $V \xrightarrow{\sim} V^*$ given by $v \mapsto \langle v, \cdot \rangle$.

Now suppose that we give an isomorphism $\phi \colon V \to W$ between two finite-dimensional vector spaces, with a strictly convex norm on each. We wish to find the operator norm $\|\phi\|_{V,W}$ by finding a nonzero vector $v \in V$ that maximizes the ratio of norms

$$\mathrm{NR}(v) := \frac{\|\phi(v)\|_W}{\|v\|_V}.$$

ALGORITHM 6.16. To attempt to maximize $\mathrm{NR}(v)$, pick $v_0 \in V$ and set $i = 0$.

(1) Let $w_i \in W$ be $\phi(v_i)$.

(2) Find the supporting vector $w_i^* \in W^*$ for w_i.

(3) Let $v_i^* \in V^*$ be $\phi^*(w_i^*)$.

(4) Find the supporting vector $v_{i+1} \in V$ for v_i^*, increase i by 1, and return to Step (1).

This gives sequences of vectors $v_i \in V$, $v_i^* \in V^*$, $w_i \in W$, and $w_i^* \in W^*$. We may also consider the corresponding sequences $[v_i]$, etc., in the respective projective spaces. The candidate for maximizing NR on PV is $\lim_{i \to \infty}[v_i]$, if it exists.

Let $\mathrm{Iter}_\phi \colon V \circlearrowleft$ be the composition of the steps in Algorithm 6.16, and let $P\,\mathrm{Iter}_\phi \colon PV \circlearrowleft$ be the corresponding map on projective spaces.

LEMMA 6.17. *NR(v) does not decrease under* Iter_ϕ: $NR(v) \leqslant NR(\mathrm{Iter}_\phi(v))$. *If we have equality, then* $[v]$ *is a fixed point of* $P\,\mathrm{Iter}_\phi$.

Proof. We use notation from Algorithm 6.16. Repeatedly apply Equation (6.14), as an equality for vectors that support each other:

$$\|w_i\|\,\|w_i^*\| = \langle w_i^*, w_i \rangle = \langle w_i^*, \phi v_i \rangle = \langle \phi^* w_i^*, v_i \rangle = \langle v_i^*, v_i \rangle$$

$$\|v_i^*\|\,\|v_{i+1}\| = \langle v_i^*, v_{i+1} \rangle = \langle \phi^* w_i^*, v_{i+1} \rangle = \langle w_i^*, \phi v_{i+1} \rangle = \langle w_i^*, w_{i+1} \rangle$$

$$\frac{\mathrm{NR}(v_i)}{\mathrm{NR}(v_{i+1})} = \frac{\|w_i\|\,\|v_{i+1}\|}{\|v_i\|\,\|w_{i+1}\|} = \frac{\langle v_i^*, v_i \rangle}{\|v_i\|\,\|w_i^*\|} \frac{\langle w_i^*, w_{i+1} \rangle}{\|w_{i+1}\|\,\|v_i^*\|}$$

$$\leqslant \frac{\|v_i^*\|\,\|v_i\|}{\|v_i\|\,\|w_i^*\|} \frac{\|w_i^*\|\,\|w_{i+1}\|}{\|w_{i+1}\|\,\|v_i^*\|} = 1.$$
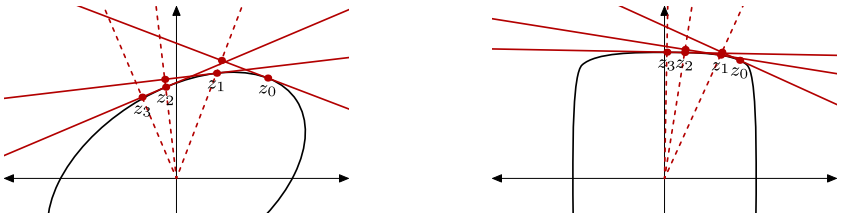
Figure 5. Examples for iteration on vector spaces (up to scale). On the left, B_1 comes from a quadratic norm as in Example 6.19. The right shows an example where Algorithm 6.16 does not converge to a global maximum of NR.

If we have equality, then $\langle v_i^*, v_i \rangle = \|v_i^*\| \|v_i\|$ and so v_i is a multiple of the supporting vector for v_i^*, namely v_{i+1}.   □

COROLLARY 6.18. *A vector $[v_0] \in PV$ that maximizes $NR(v_0)$ is a fixed point for $P\,\mathrm{Iter}_\phi$. If v_0 is an attracting fixed point for the $P\,\mathrm{Iter}_\phi$, then $\|\phi v\|/\|v\|$ has a local maximum at v_0.*

EXAMPLE 6.19. In the setting of Example 6.15, if the norms on V and W come from inner products, then Iter_ϕ is $\phi^*\phi$ and the iteration reduces to power iteration: find the maximum eigenvector of $\phi^*\phi$ by repeatedly applying it. This almost always converges to an eigenvector of maximal eigenvalue, with convergence rate determined by the ratio between the two largest distinct eigenvalues.

EXAMPLE 6.20. Consider the case when ϕ is the identity on $\mathbb{R}^n$ and $\|\cdot\|_1$ is the standard inner product. Then $\|\cdot\|_2$ is defined by its unit norm ball $B_2 \subset \mathbb{R}^n$. The supporting vector of $v \in \partial B_1$ is the tangent hyperplane to B_2. Up to scale, Algorithm 6.16 alternates between taking a tangent hyperplane to B_2, finding the closest point to the origin on the tangent hyperplane, and projecting to B_2, as in Figure 5.

We now return to the actual case of interest of elastic graphs. For a marked elastic graph $G = (\Gamma, \alpha)$, we think of a 'duality' between homotopy classes of maps $f : G \to K$ to a length graph and maps $c : W \to G$ from a weighted graph. The pairing is given by $\langle [c], [f] \rangle = \ell[f \circ c]$, and the two norms are $\sqrt{\mathrm{Dir}[f]}$ and $\sqrt{\mathrm{EL}[c]}$. The analogue of Equation (6.14) is Equation (2.20). More concretely, we work with $\mathcal{L}(\Gamma)$ and $\mathcal{W}(\Gamma)$.

DEFINITION 6.21. For G an elastic graph, the *pairing* $\langle \cdot, \cdot \rangle \colon \mathcal{W}(\Gamma) \times \mathcal{L}(\Gamma) \to \mathbb{R}$

is defined by

$$\langle w, \ell \rangle := \sum_{e \in \mathrm{Edge}(\Gamma)} w(e)\ell(e)$$

and the *duality map* $D_G \colon \mathcal{W}(\Gamma) \to \mathcal{L}(\Gamma)$ is defined by

$$D_G(w)(e) := \alpha(e)w(e).$$

For $w \in \mathcal{W}(\Gamma)$ a balanced set of weights and $\ell \in \mathcal{L}(\Gamma)$ any set of lengths, $\langle w, \ell \rangle = \ell[f \circ c]$ where $c \colon (\Gamma, w) \to G$ and $f \colon G \to (\Gamma, \ell)$ are both the identity map.

Fix $[\phi] \colon G_1 \to G_2$ a homotopy class of maps between marked elastic graphs $G_1 = (\Gamma_1, \alpha_1)$ and $G_2 = (\Gamma_2, \alpha_2)$. For $\ell \in \mathcal{L}(\Gamma_2)$, let $\mathrm{DR}(\ell)$ be the ratio of Dirichlet energies $\mathrm{Dir}_\ell^{\alpha_1}[\phi]/\mathrm{Dir}_\ell^{\alpha_2}[\mathrm{id}_{\Gamma_2}]$, and for $w \in \mathcal{W}(\Gamma_1)$, let $\mathrm{ER}(w)$ be the ratio of extremal lengths $\mathrm{EL}_{\alpha_2}^w[\phi]/\mathrm{EL}_{\alpha_1}^w[\mathrm{id}_{\Gamma_1}]$.

ALGORITHM 6.22. To attempt to maximize DR and ER, pick a generic initial set of lengths $\ell_0 \in \mathcal{L}(\Gamma_2)$, and let K_0 be the marked length graph (Γ_2, ℓ_0). We will write D_i for D_{G_i}. Let $f_0 \colon G_2 \to K_0$ be the identity on the level of graphs. Set $i = 0$ and consider the following iteration.

(1) Find a harmonic representative g_i of the composite map $[\phi \circ f_i] \colon G_1 \to K_i$. For each edge e of Γ_1, let $m_i(e) = \ell(g_i(e))$. Thus $m_i = H_{[f]}(\ell_i)$, with $H_{[f]}$ (Definition 5.18).

(2) Set $w_i = D_1^{-1}(m_i) \in \mathcal{W}(\Gamma_1)$, so that $w_i(e)$ is $|g_i'(e)|$. Let W_i be the tension-weighted graph (Γ_1, w_i), and let $d_i \colon W_i \to G_1$ be the identity on the level of graphs. W_i is balanced by Proposition 5.2.

(3) Set $v_i = N_{[\phi]}(w_i)$, the push-forward of w_i. Since g_i is harmonic and thus taut from the tension-weighted graph, we have

$$v_i(y) = \sum_{x \in g_i^{-1}(y)} w_i(x). \tag{6.23}$$

The weights v_i are also balanced, by Proposition 3.18. Let $c_i \colon W_i \to G_2$ be $g_i \circ d_i$.

(4) Set $\ell_{i+1} = D_2(v_i) \in \mathcal{L}(\Gamma_2)$, let $K_{i+1} = (\Gamma_2, \ell_{i+1})$, and let $f_{i+1} \colon G_2 \to K_{i+1}$ be the identity on the level of graphs. Increase i by 1 and return to Step (1).

Schematically, Algorithm 6.22 iterates around the following loop.

$$
\begin{array}{ccc}
\ell_i \in \mathcal{L}(\Gamma_2) & \xleftarrow{\quad D_2 \quad} & \mathcal{W}(\Gamma_2) \ni v_i \\
{\scriptstyle H_{[\phi]}} \downarrow & & \uparrow {\scriptstyle N_{[\phi]}} \\
m_i \in \mathcal{L}(\Gamma_1) & \xrightarrow{\quad D_1^{-1} \quad} & \mathcal{W}(\Gamma_1) \ni w_i.
\end{array}
\tag{6.24}
$$

If all lengths and weights remain positive, we have a diagram of spaces and maps, in which:

- rows are tight sequences (see below);

- the dashed lines are only defined up to homotopy; and

- the regions commute up to homotopy:

$$
\begin{array}{ccccc}
W_{i-1} & \xrightarrow{\quad c_{i-1} \quad} & G_2 & \xrightarrow{\quad f_i \quad} & K_i \\
& {\scriptstyle [\phi]} \nearrow & & & \| \\
W_i & \xrightarrow{\ d_i\ } G_1 & \xrightarrow{\quad g_i \quad} & & K_i \\
\| & & {\scriptstyle [\phi]} \searrow & & \\
W_i & \xrightarrow{\quad c_i \quad} & G_2 & \xrightarrow{\ f_{i+1}\ } & K_{i+1}.
\end{array}
\tag{6.25}
$$

The row $W_i \xrightarrow{d_i} G_1 \xrightarrow{g_i} K_i$ is tight since g_i is harmonic and $w_i = |g_i'|$, while the row $W_i \xrightarrow{c_i} G_2 \xrightarrow{f_{i+1}} K_{i+1}$ is tight because c_i is taut and $\ell_{i+1} = D_2(v_i)$.

Let $\mathrm{Iter}_\phi \colon \mathcal{L}(\Gamma_2) \circlearrowleft$ be the composition

$$
\mathrm{Iter}_\phi = D_2 \circ N_{[\phi]} \circ D_1^{-1} \circ H_{[\phi]}.
$$

All of the maps involved are piecewise-linear and linear on rays by Propositions 3.44 and 5.19, so Iter_ϕ is as well. We allow some of the lengths or weights to vanish (that is, we include the boundary of $\mathcal{L}(\Gamma_i)$ and $\mathcal{W}(\Gamma_i)$ in the maps).

LEMMA 6.26. *If $\phi \colon G_1 \to G_2$ is an essentially surjective map between marked elastic graphs and $0 \neq \ell \in \mathcal{L}(\Gamma_2)$, then $\mathrm{Iter}_\phi(\ell) \neq 0$.*

Proof. Let e_2 be an edge of G_2 with $\ell(e_2) \neq 0$ and let $f \colon G_1 \to (\Gamma_2, \ell)$ be the harmonic representative. Since ϕ is essentially surjective, there must be an edge e_1 of G_1 with so that f is not constant on e_1 and $f(e_1)$ intersects e_2. Set $v = N_{[\phi]}(D_1^{-1}(H_{[\phi]}(\ell)))$ as in Algorithm 6.22. By Equation (6.23), $v(e_2) \geqslant |f'(e_1)| > 0$. □

As a result of linearity on rays and Lemma 6.26, Iter_ϕ descends to a projective map $P\,\text{Iter}_\phi\colon P\mathcal{L}(\Gamma_2)\circlearrowleft$. Since $P\mathcal{L}(\Gamma_2)$ is a compact ball, there is a fixed point.

The iteration is parallel to the iteration on vector spaces, with

$$
\begin{aligned}
V &\leftrightarrow \mathcal{L}(G_2) & W &\leftrightarrow \mathcal{L}(G_1)\\
v_i &\leftrightarrow f_i \text{ or } \ell_i & w_i &\leftrightarrow g_i \text{ or } m_i\\
v_i^* &\leftrightarrow c_i \text{ or } v_i & w_i^* &\leftrightarrow d_i \text{ or } w_i.
\end{aligned}
$$

REMARK 6.27. The only computationally expensive step in Algorithm 6.22 is Step (1), finding the harmonic equilibrium.

LEMMA 6.28. *Let* $[\phi]\colon G_1 \to G_2$ *be an essentially surjective homotopy class of maps between marked elastic graphs.* Iter_ϕ *does not decrease the objective functions for both Dirichlet energy and extremal length: with the notation from Algorithm 6.22, for* $i \geqslant 0$,

$$
DR(\ell_i) \leqslant DR(\ell_{i+1}) \qquad\qquad ER(w_i) \leqslant ER(w_{i+1}).
$$

If either inequality is an equality, then $[\ell_i]$ *is a fixed point of* $P\,\text{Iter}_\phi$.

Proof. This is parallel to Lemma 6.17. Tightness of the rows in Diagram (6.25) tells us

$$
\text{Dir}[g_i]\,\text{EL}[d_i] = \ell[g_i \circ d_i]^2 = \ell[f_i \circ \phi \circ d_i]^2 = \ell[f_i \circ c_i]^2
$$
$$
\text{Dir}[f_{i+1}]\,\text{EL}[c_i] = \ell[f_{i+1} \circ c_i]^2 = \ell[f_{i+1} \circ \phi \circ d_i]^2 = \ell[g_{i+1} \circ d_i]^2
$$

$$
\begin{aligned}
\frac{DR(\ell_i)}{DR(\ell_{i+1})} &= \frac{\text{Dir}[g_i]}{\text{Dir}[f_i]}\frac{\text{Dir}[f_{i+1}]}{\text{Dir}[g_{i+1}]} = \frac{\ell[f_i \circ c_i]^2}{\text{Dir}[f_i]\,\text{EL}[d_i]}\frac{\ell[g_{i+1} \circ d_i]^2}{\text{Dir}[g_{i+1}]\,\text{EL}[c_i]}\\
&\leqslant \frac{\text{Dir}[f_i]\,\text{EL}[c_i]}{\text{Dir}[f_i]\,\text{EL}[d_i]}\frac{\text{Dir}[g_{i+1}]\,\text{EL}[d_i]}{\text{Dir}[g_{i+1}]\,\text{EL}[c_i]} = 1.
\end{aligned}
$$

If we have equality, then $\ell[f_i \circ c_i]^2 = \text{Dir}[f_i]\,\text{EL}[c_i]$ and so $W_i \xrightarrow{c_i} G_2 \xrightarrow{f_i} K_i$ is a tight sequence, in addition to $W_i \xrightarrow{c_i} G_2 \xrightarrow{f_{i+1}} K_{i+1}$. In a tight sequence $W \xrightarrow{c} G \xrightarrow{f} K$, Equation (2.20) is an equality. The Cauchy–Schwarz inequality used in its proof then implies that $|f'|$ is proportional to n_c, which in the present case says that ℓ_i and ℓ_{i+1} must be multiples of each other, as desired for the last statement.

The proof for the inequality on ER is parallel:

$$
\frac{ER(w_i)}{ER(w_{i+1})} = \frac{\text{EL}[c_i]}{\text{EL}[d_i]}\frac{\text{EL}[d_{i+1}]}{\text{EL}[c_{i+1}]} = \frac{\ell[g_{i+1} \circ d_i]^2}{\text{EL}[d_i]\,\text{Dir}[f_{i+1}]}\frac{\ell[f_{i+1} \circ c_{i+1}]}{\text{EL}[c_{i+1}]\,\text{Dir}[g_{i+1}]} \leqslant 1. \qquad \square
$$

Now consider the lengths in $[\ell] \in P\mathcal{L}(\Gamma_2)$ that maximize $DR(\ell)$. By Lemma 6.28, $[\ell]$ is a fixed point for $P\operatorname{Iter}_\phi$. We have a partial converse for interior fixed points.

PROPOSITION 6.29. *Let* $[\phi]\colon G_1 \to G_2$ *be an essentially surjective homotopy class of maps between marked elastic graphs and suppose* $[\ell] \in P\mathcal{L}^+(\Gamma_2)$ *is an interior fixed point of* $P\operatorname{Iter}_\phi$ *with multiplier* λ. *Let* $K = (\Gamma_2, \ell)$, *let* $f\colon G_2 \to K$ *be the identity on the level of graphs, and let* $g\colon G_1 \to K$ *be a harmonic representative of* $\phi \circ f$. *Then* $f^{-1} \circ g\colon G_1 \to G_2$ *is* λ-*filling.*

Proof. At the fixed point, the tight sequences from Diagram (6.25) collapse (up to scale) to

$$
\begin{array}{ccc}
 & \xrightarrow{\ c\ } G_2 \xrightarrow{\ f\ } & \\
W & \overset{[\phi]}{\dashrightarrow} & K \\
 & \searrow_{d} \quad \quad \quad _{g} \nearrow & \\
 & G_1 &
\end{array}
\tag{6.30}
$$

where the top and bottom are tight sequences and the two triangles commute up to homotopy. By hypothesis, K is a positive length graph and f is invertible. Set $\psi = f^{-1} \circ g$, so $\psi \in [\phi]$. We will show that ψ is λ-filling.

We have chosen $\ell \in \mathcal{L}(\Gamma_2)$. Choose m, w, and v so that

$$
\begin{aligned}
m &= H_{[\phi]}(\ell) \in \mathcal{L}(\Gamma_1) \\
w &= D_1^{-1}(m) \in \mathcal{W}(\Gamma_1) \\
\lambda v &= N_{[\phi]}(w) \in \mathcal{W}(\Gamma_2) \\
\ell &= D_2(v) \in \mathcal{L}(\Gamma_2).
\end{aligned}
\tag{6.31}
$$

The weights w and v are balanced as in Algorithm 6.22, so we have balanced strip graphs $S_1 = (\Gamma_1, w, m)$ and $S_2 = (\Gamma_2, v, \ell)$ compatible with α_1 and α_2. The definition of $H_{[\phi]}$ ensures that ψ preserves lengths, and the fact that ψ is taut ensures that it scales weights by a factor of λ. $\square$

COROLLARY 6.32. *Any fixed point* $[\ell]$ *for* $P\operatorname{Iter}_\phi$ *in the interior of* $P\mathcal{L}(\Gamma_2)$ *is a global maximum for* $DR(\ell)$.

Proof. Immediate from Propositions 6.29 and 6.8. $\square$

REMARK 6.33. The maps in Iter_ϕ can be given an interpretation in terms of derivatives. Specifically, let $\operatorname{Dir}_{[\phi]}\colon \mathcal{L}(\Gamma_2) \to \mathbb{R}$ be Dirichlet energy as a function of lengths from Proposition 5.19. Then $\operatorname{Dir}_{[\phi]}$ is C^1 with derivative given by

$$
d\operatorname{Dir}_{[\phi]}(\ell) = 2 \cdot N_{[\phi]}(D_{G_1}^{-1}(H_{[\phi]}(\ell))) \in \mathcal{W}(\Gamma_2) = \mathcal{L}(\Gamma_2)^*.
$$

Recall from the introduction that, physically, tension in a spring is the derivative of energy as the length varies. Thus $d \operatorname{Dir}_{[\phi]}(\ell)$ gives the total tension in each edge of Γ_2.

6.3.   Behavior on the boundary.   We now turn to boundary fixed points of $P \operatorname{Iter}_\phi$. Boundary fixed points need not be a global maxima of DR.

EXAMPLE 6.34.   In Example 6.9, $P\mathcal{L}(\Gamma_1)$ is an interval, and both endpoints of the interval are fixed points for $P \operatorname{Iter}_\phi$. One endpoint is attracting (and the maximum of DR) and the other is repelling (and the minimum of DR).

However, we can still extract much useful structure from a boundary fixed point.

PROPOSITION 6.35.   *Let* $[\phi] \colon (\Gamma_1, \alpha_1) \to (\Gamma_2, \alpha_2)$ *be an essentially surjective homotopy class of maps between marked elastic graphs, and suppose that* $[\ell] \in \partial P\mathcal{L}(\Gamma_2)$ *is a fixed point of* $P \operatorname{Iter}_\phi$ *with multiplier* $\lambda > 0$. *Then there are decompositions of* Γ_i *into complementary subgraphs* $\Gamma_i = \Delta_i \cup \Sigma_i$, *along with a map* $\psi \colon \Gamma_1 \to \Gamma_2$ *so that:*

(1) $\psi \in [\phi]$;

(2) *the edges of* Σ_2 *are those edges* e *with* $\ell(e) = 0$;

(3) $\psi(\Delta_1) \subset \Delta_2$ *and* $\psi(\Sigma_1) \subset \Sigma_2$; *and*

(4) *the restriction of* ψ *to a map from* (Δ_1, α_1) *to* (Δ_2, α_2) *is* λ-*filling.*

The key thing missing from Proposition 6.35 in asserting that ψ is partially λ-filling is condition (6) of Definition 6.10. (Indeed Example 6.9, shows that we cannot get a partially λ-filling map from an arbitrary boundary fixed point.)

Proof. Define m, w, and v from ℓ by Equation (6.31). There are tight sequences similar to Diagram (6.30), with *weak* harmonic maps $(\widetilde{f}, f)$ and $(\widetilde{g}, g)$:

$$
\begin{array}{ccccc}
 & & \Gamma_2 & \xrightarrow{\;\widetilde{f}\;} & K \\
 & \nearrow^{c} & & & \\
W & & \widetilde{g} & & \downarrow \kappa \\
 & \searrow_{d} & & & \\
 & & \Gamma_1 & \xrightarrow[\;g\;]{} & K^*.
\end{array}
\qquad (6.36)
$$

Choose $\widetilde{f}$ to be the identity (since Γ_2 and K have the same underlying graph). Choose $\widetilde{g}$ so that it is vertex-precise and $\widetilde{g} \circ d$ is taut, as guaranteed by Proposition 5.8.

Let $\Sigma_2 \subset \Gamma_2$ be the null subgraph of K. Let Δ_2 be the complementary subgraph. Similarly let Σ_1 be the subgraph of Γ_1 whose edges are those on which w and m are 0 and let Δ_1 be the complementary subgraph. Set $\psi = \widetilde{f}^{-1} \circ \widetilde{g}$; we must check ψ satisfies the conditions.

The first two conditions are immediate. To check condition (3), pick a regular value y for ψ on an edge e_2 of Γ_2 and a preimage x on an edge e_1 of Γ_1. We must show that $y \in \Sigma_2 \Leftrightarrow x \in \Sigma_1$. If $y \in \Sigma_2$, then since $\widetilde{g} \circ d$ is taut, we have $n_{\widetilde{g}}^w(y) = n_{[\phi]}^w(y) = \lambda v(e_2) = 0$. We therefore must have $w(e_1) = 0$, so $x \in \Sigma_1$. Conversely, if $x \in \Sigma_1$, then by definition $\ell(e_1) = 0$. Since g is length-preserving as a map from (Γ_1, ℓ) to K^* and the lift $\widetilde{g}$ is vertex-precise, $\widetilde{g}(e_1)$ lies inside the null subgraph of K (since by assumption $\widetilde{g}$ is not constant on e_1), so $y \in \Sigma_2$.

Let $\psi|_\Delta$ be the restriction of ψ to a map from Δ_1 to Δ_2. It remains to show that

$$\psi|_\Delta \colon (\Delta_1, w|_{\Delta_1}, m|_{\Delta_1}) \longrightarrow (\Delta_2, v|_{\Delta_2}, \ell|_{\Delta_2})$$

is λ-filling as a map between strip graphs. As in Proposition 6.29, the definition of m as $m(e) = \ell(g(e)) = \ell(\psi(e))$ ensures that ψ is length-preserving. We chose $\widetilde{g} \circ d$ to be taut, and by restricting first the domain to Δ_1 by Lemma 3.40 and then the range to Δ_2 by Lemma 3.43, we see that $\psi|_\Delta$ is taut. The definition of v then tells us that $\psi|_\Delta$ scales weights by λ. □

In the setting of Proposition 6.35, let $\psi|_\Sigma$ be the restriction of ψ to a map between the collapsed subgraphs Σ_1 and Σ_2. For $i = 1, 2$, let $P_i' = \Sigma_i \cap \Delta_i$ be the vertices shared between the two complementary subgraphs, and let P_i be the union of P_i' with the vertices of Σ_i that were already marked. We view $\psi|_\Sigma$ as a map of marked elastic graphs $(\Sigma_1, P_1, \alpha_1) \to (\Sigma_2, P_2, \alpha_2)$, and consider the problem of finding $\text{Emb}[\psi|_\Sigma]$ in its own right.

PROPOSITION 6.37. *In the setting of Proposition 6.35, suppose that there is a fixed point $[\ell_0]$ of $P\,\text{Iter}_{\psi|_\Sigma}$ with multiplier $\lambda_0 > \lambda$. Then ℓ is not a local maximum of $DR(\ell)$.*

Proof. We will show that $DR(\ell + \varepsilon\ell_0) > DR(\ell)$ for sufficiently small ε. From $\ell_0 \in \mathcal{L}(\Sigma_2)$, construct $m_0 \in \mathcal{L}(\Sigma_1)$, $w_0 \in \mathcal{W}(\Sigma_1)$, and $v_0 \in \mathcal{W}(\Sigma_2)$ by Equation (6.31). Extend ℓ_0, m_0, w_0, and v_0 to Γ_i by setting them to be zero on edges in Δ_i. Let $\psi_0 \colon (\Sigma_1, P_1) \to (\Sigma_2, P_2)$ be the map constructed from ℓ_0 by Proposition 6.35. In particular, ψ_0 is harmonic as a map from $(\Sigma_1, P_1, \alpha_1)$ to (Σ_2, P_2, ℓ_0). Define a new map $h \colon \Gamma_1 \to \Gamma_2$ by

$$h(x) := \begin{cases} \psi(x) & x \in \Delta_1, \\ \psi_0(x) & x \in \Sigma_1. \end{cases} \tag{6.38}$$

Since we pinned $\Delta_i \cap \Sigma_i$ in $\psi|_\Sigma$, the map h is continuous. By construction, h is weakly harmonic as a map from (Γ_1, α_1) to (Γ_2, ℓ). We claim that h is also harmonic if we perturb ℓ. For small ε, consider the modified lengths $\ell_\varepsilon = \ell + \varepsilon\ell_0 \in \mathcal{L}(\Gamma_2)$. Let h_ε be h considered as a map from (Γ_1, α_1) to $(\Gamma_2, \ell_\varepsilon)$.

CLAIM 6.39. *For ε sufficiently small, h_ε is a harmonic map.*

Proof. For simplicity, we suppose that ℓ_0 is nonzero on every edge, so that $(\Gamma_2, \ell_\varepsilon)$ is a length graph and h_ε is an ordinary map (not weak). The case when ℓ_0 has some zeros can be treated by induction.

The tension weight of h is $w \in \mathcal{W}(\Gamma_1)$. Let w_ε be the tension weight of h_ε. Concretely,

$$w_\varepsilon(e) = \begin{cases} w(e) & e \in \mathrm{Edge}(\Delta_1), \\ \varepsilon w_0(e) & e \in \mathrm{Edge}(\Sigma_1). \end{cases}$$

We must check that h_ε is still taut as a map from $(\Gamma_1, w_\varepsilon)$ to Γ_2. By Proposition 5.2, we must check that w_ε satisfies the train-track triangle inequalities at the vertices of Γ_1.

For vertices in $\Delta_1 \backslash \Sigma_1$ and $\Sigma_1 \backslash \Delta_1$, the triangle inequalities follow from the fact that ψ and ψ_0 are harmonic, respectively. For vertices in $\Delta_1 \cap \Sigma_1$, the triangle inequalities follow from Lemma 6.40, where the a_i are the weights of the incident groups of edges of Δ_1 and the b_i are the weights of the incident groups of edges of Σ_1. □

LEMMA 6.40. *If $(a_1, \ldots, a_n) \in \mathbb{R}_+^n$ satisfies the triangle inequalities and $(b_1, \ldots, b_m) \in \mathbb{R}_+^m$ is another vector, then, for all ε sufficiently small,*

$$(a_1, \ldots, a_n, \varepsilon b_1, \ldots, \varepsilon b_m)$$

satisfies the triangle inequalities.

Proof. Elementary. □

Returning to the proof of Proposition 6.37, since h_ε is harmonic, if f and g are the harmonic maps to (Γ_2, ℓ) and f_0 and g_0 are the harmonic maps to (Σ_2, ℓ_0), we have

$$\mathrm{DR}(\ell_\varepsilon) = \frac{\mathrm{Dir}(h_\varepsilon)}{\mathrm{Dir}(f) + \varepsilon^2 \, \mathrm{Dir}(f_0)} = \frac{\mathrm{Dir}(g) + \varepsilon^2 \, \mathrm{Dir}(g_0)}{\mathrm{Dir}(f) + \varepsilon^2 \, \mathrm{Dir}(f_0)} > \frac{\mathrm{Dir}(g)}{\mathrm{Dir}(f)} = \lambda = \mathrm{DR}(\ell)$$

using the assumption that $\lambda_0 = \mathrm{Dir}(g_0)/\mathrm{Dir}(f_0) > \lambda$. □

We can now prove the promised existence of a partially λ-filling map.

Proof of Proposition 6.12. First, if $[\phi]$ is not essentially surjective, we can restrict to the essential image of $[\phi]$, the image of a taut representative of $[\phi]$ with respect to any weight structure on G_1. After doing this, $P \operatorname{Iter}_\phi$ is defined by Lemma 6.26.

We proceed by induction on the size of Γ_1. Given the homotopy class $[\phi]$, let $\ell \in \mathcal{L}(\Gamma_1)$ be a global maximum of DR (which exists since $P\mathcal{L}(\Gamma_1)$ is compact). By Lemma 6.28, ℓ is a fixed point of $P \operatorname{Iter}_\phi$; let λ be its multiplier. If ℓ is in $\mathcal{L}^+(\Gamma_1)$, we are done by Proposition 6.29. Otherwise, consider the subgraphs Δ_i and Σ_i given by Proposition 6.35, with restricted maps $\psi|_\Delta$ (which is λ-filling) and $\phi|_\Sigma$. Since Σ_1 is a proper subgraph of Γ_1, by induction we can find a partially λ_1-filling map $\psi|_\Sigma \in [\phi_\Sigma]$ for some $\lambda_1 \geqslant 0$.

Now assemble $\psi|_\Delta$ and $\psi|_\Sigma$ to a single map $\psi \colon \Gamma_1 \to \Gamma_2$ by Equation (6.38). Then ψ is λ-filling on Δ_1 and partially λ_1-filling on Σ_1. By Proposition 6.37, $\lambda_1 \leqslant \lambda$, so ψ is partially λ-filling on all of Γ_1. $\qquad\square$

REMARK 6.41. The map constructed above has a stronger 'layered' structure, where Γ_1 and Γ_2 are divided into layers, each with its own filling constant. Specifically, there are properly nested subgraphs

$$\Gamma_1 = \Sigma_1^0 \supsetneq \Sigma_1^1 \supsetneq \cdots \supsetneq \Sigma_1^n$$
$$\Gamma_2 = \Sigma_2^0 \supsetneq \Sigma_2^1 \supsetneq \cdots \supsetneq \Sigma_2^n,$$

a map $\psi \colon \Gamma_1 \to \Gamma_2$ in $[\phi]$, and a sequence of filling constants $\lambda_1 > \cdots > \lambda_n$, with the following properties.

(1) ψ preserves Σ^i: for $1 \leqslant i \leqslant n$, $\psi(\Sigma_1^i) \subset \Sigma_2^i$.

(2) ψ preserves $\Sigma^{i-1} \backslash \Sigma^i$: for $k \in \{1, 2\}$ and $0 < i \leqslant n$, let Δ_k^i be the complementary subgraph to Σ_k^i in Σ_k^{i-1}. Then $\psi(\Delta_1^i) \subset \Delta_2^i$.

(3) ψ is λ_i-filling on Δ^i: for $0 \leqslant i \leqslant n$, let ψ_i be the restriction of ψ to a map from Δ_1^i to Δ_2^i, where for $i > 0$ we additionally mark the vertices in $\Delta_1^i \cap \Delta_1^{i-1}$ and in $\Delta_2^i \cap \Delta_2^{i-1}$. Then ψ_i is λ_i-filling.

(4) $\psi|_{\Sigma_1^n}$ is constant.

6.4. General targets.

We turn to Theorem 2, allowing more general length space targets. First we need to generalize Equation (2.23) to this setting.

LEMMA 6.42. *Let G_1 and G_2 be elastic graphs, let X be a length space, let $\phi \colon G_1 \to G_2$ be a PL map, and let $f \colon G_2 \to X$ be a Lipschitz map. Then*

$$\operatorname{Dir}(f \circ \phi) \leqslant \operatorname{Emb}(\phi) \operatorname{Dir}(f).$$

Proof. We compute

$$\mathrm{Dir}(f \circ \phi) = \int_{x \in G_1} |(f \circ \phi)'(x)|^2 \, dx$$

$$= \int_{y \in G_2} \left(\sum_{x \in \phi^{-1}(y)} |\phi'(x)| \right) |f'(y)|^2 \, dy$$

$$\leqslant \mathrm{Emb}(\phi) \, \mathrm{Dir}(f).$$

In the second line we do a change of variables from G_1 to G_2, using $dx = |\phi'(x)| \, dy$. (Any portions of G_1 where ϕ is constant and so $\phi^{-1}(y)$ is uncountable do not contribute to the integrals.) In the last line we use $\int |a| \cdot |b| \, dy \leqslant \mathrm{ess} \sup |a| \cdot \int |b| \, dy$. □

Proof of Theorem 2. Suppose $\mathrm{Emb}(\phi)$ is minimal within the homotopy class $[\phi]$ and that $\mathrm{Dir}(f)$ is within a multiplicative factor of ε of the infimum. Then

$$\mathrm{Dir}[f \circ \phi] \leqslant \mathrm{Dir}(f \circ \phi) \leqslant \mathrm{Emb}(\phi) \, \mathrm{Dir}(f) \leqslant \mathrm{Emb}[\phi] \, \mathrm{Dir}[f](1 + \varepsilon).$$

Since we can choose ε as small as we like, this gives one inequality of the desired equality. The other direction comes from Theorem 1. □

6.5. Algorithmic questions. Given a homotopy class $[\phi] \colon G_1 \to G_2$ of maps between marked elastic graphs, we have proved that Iter_ϕ has a projectively fixed set of lengths $\ell \in \mathcal{L}(G_2)$ maximizing $\mathrm{DR}(\ell)$ and giving a partially λ-filling map in $[\phi]$. The lengths maximizing $\mathrm{DR}(\ell)$ need not be unique.

EXAMPLE 6.43. Let G_1 and G_2 both be the join of two circles, as in Example 6.9, with all elastic constants equal to 1, and let ϕ be the identity map. Then Iter_ϕ is the identity and DR is constant on $\mathcal{L}(G_2)$.

QUESTION 6.44. What is the structure of the subset of $\mathcal{L}(G_2)$ on which DR reaches its maximum value? For instance, is it a convex subset of $\mathcal{L}(G_2)$? Can it be a proper subset of the interior of $\mathcal{L}(G_2)$?

As for convergence, we would like to say that Algorithm 6.22 works, in the sense that iterating $P \, \mathrm{Iter}_\phi$ always converges to a maximum of DR (which also computes $\mathrm{Emb}[\phi]$). The presence of extra fixed points of Iter_ϕ means that this does not always happen. (In dynamical terms, DR is not decreased by Iter_ϕ, and not necessarily increased; so DR is not quite a Lyapunov function for this discrete dynamical system.) But we can make some statements.

PROPOSITION 6.45. *Algorithm 6.22 gives a sequence of lengths $\ell_i \in \mathcal{L}(G_2)$ that converge projectively to a fixed point for $P \operatorname{Iter}_\phi$.*

Proof. By Lemma 6.28, $\operatorname{DR}(\ell_i)$ does not decrease and has an upper bound; thus $\operatorname{DR}(\ell_i)$ has a limit, and the $[\ell_i]$ have an accumulation point $[\ell_\infty]$ with $\operatorname{DR}(\operatorname{Iter}_\phi(\ell_\infty)) = \operatorname{DR}(\ell_\infty)$. By Lemma 6.28 again, $[\ell_\infty]$ is a fixed point of $P \operatorname{Iter}_\phi$, and therefore the $[\ell_i]$ limit to $[\ell_\infty]$, without the need to pass to a subsequence. $\square$

LEMMA 6.46. *For $[\phi]\colon G_1 \to G_2$ a homotopy class of maps between marked elastic graphs, the set $\{DR(\ell) \mid \ell$ a fixed point of $P \operatorname{Iter}_\phi\}$ is finite.*

Proof. Let $\ell \in \mathcal{L}(G_2)$ be a projective fixed point for Iter_ϕ with multiplier λ. Proposition 6.35 gives a λ-filling map on subgraphs $\phi|_\Delta\colon \Delta_1 \to \Delta_2$. By Proposition 6.8, $\lambda = \operatorname{Emb}[\phi|_\Delta]$ and thus depends only on the subgraphs Δ_i. Since there are only finitely many subgraphs, we are done. $\square$

PROPOSITION 6.47. *For $\phi\colon G_1 \to G_2$ a homotopy class of maps between marked elastic graphs, there is an open subset of $P\mathcal{L}(G_2)$ on which $P \operatorname{Iter}_\phi$ converges to a maximum of DR.*

Proof. Let λ be the maximum value of DR on $P\mathcal{L}(G_2)$. By Lemma 6.46, there is an $\varepsilon > 0$ so that there are no fixed points of $P \operatorname{Iter}_\phi$ in $\operatorname{DR}^{-1}(\lambda - \varepsilon, \lambda)$. Then by Proposition 6.45, $P \operatorname{Iter}_\phi$ converges to a maximum of DR on $\operatorname{DR}^{-1}(\lambda - \varepsilon, \lambda]$. $\square$

QUESTION 6.48. Does Algorithm 6.22 converge to a maximum of DR for an open dense set of initial data?

A few words are in order on why Question 6.48 is not as easy as it may appear. If $[\ell_1]$ is a fixed point of $P \operatorname{Iter}_\phi$ with $\operatorname{DR}(\ell_1) < \operatorname{Emb}[\phi]$, then by Proposition 6.37 it is not a local maximum of DR. If Iter_ϕ were linear, that would imply the set attracted to $[\ell_1]$ has empty interior. Since Iter_ϕ is only PL, the situation is more complicated. For instance, Iter_ϕ can map an open subset of $\mathcal{L}^+(G_2)$ to a subset $S \subset \partial\mathcal{L}(G_2)$, since harmonic maps can generically map vertices to vertices. Then S could potentially be attracted to $[\ell_1]$.

We can nevertheless improve Algorithm 6.22 to always find $\operatorname{Emb}[\phi]$.

ALGORITHM 6.49. Given an essentially surjective homotopy class $[\phi]\colon G_1 \to G_2$ of maps between marked elastic graphs, to find $\operatorname{Emb}[\phi]$, pick arbitrary nonzero initial lengths $\ell_0 \in \mathcal{L}(G_2)$ and iterate Algorithm 6.22 to get a sequence of lengths ℓ_i. Since Iter_ϕ is PL, each set of lengths ℓ_i is in a closed cone of linearity

$R_i \subset \mathcal{L}(G_2)$. Since there are only finitely many domains of linearity, there must be some i and k so that $R_{i+k} = R_i$. By standard linear algebra techniques, we can see if $(\mathrm{Iter}_\phi)^k$ has a projective fixed point in R_i. By Proposition 6.45, the ℓ_i converge to a projective fixed point, so eventually the linear algebra check will succeed, giving projectively fixed lengths ℓ_∞ with multiplier λ_∞.

If ℓ_∞ is nonzero on every edge, we are done by Proposition 6.29. Otherwise, apply Proposition 6.35 to extract a map

$$\psi|_\Sigma \colon (\Sigma_1, P_1, \alpha_1) \to (\Sigma_2, P_2, \alpha_2)$$

between simpler graphs, with its own embedding energy λ_Σ (which we can find by induction). If $\lambda_\Sigma < \lambda_\infty$, we have found a partially λ_∞-filling map and are done. Otherwise, by Proposition 6.37, we can find nearby lengths ℓ_0' with $\mathrm{DR}(\ell_0') > \mathrm{DR}(\ell_\infty) \geqslant \mathrm{DR}(\ell_0)$. In this case, repeat the algorithm, with ℓ_0' in place of ℓ_0. By Lemma 6.46, eventually we will find the true maximum value of $\mathrm{DR}(\ell)$, and thus the true value of $\mathrm{Emb}[\phi]$.

In practice, Algorithm 6.49 appears to run quickly, at least for small examples. The additional steps to continue past a projective fixed point have been unnecessary. Theoretically, there is no reason to expect it to always perform well. In the closely related case of pseudo-Anosov maps, Bell and Schleimer have given examples where the analogous algorithm is slow [**BS17**].

Acknowledgments

Appendix A. General graph energies

As suggested by the notation in Definition 1.30, the energies of this paper fit into a more general framework. We start with a notion of p-conformal graphs, simultaneously generalizing weighted graphs, elastic graphs, and length graphs. There are several different perspectives. A p-conformal graph can be viewed as:

- a graph with a p-length $\alpha(e)$ on each edge;

- an equivalence class of strip graphs (Γ, w, ℓ) under a rescaling operation; or

- an equivalence class of spaces X with a length metric ℓ and measure μ under another rescaling operation.

We start with the metric view, since it is most standard, although the formulas may appear unmotivated.

DEFINITION A.1. For $p \in (1, \infty]$, a *p-conformal graph* $G^p = (\Gamma, \alpha)$ is a graph with a positive *p-length* $\alpha(e)$ on each edge e, which gives a metric. For $p = 1$, a *1-conformal graph* is a weighted graph.

The weights associated to edges in a 1-conformal graph behave differently than the p-lengths for $p > 1$; for instance, p-lengths add for two edges joined in series, while weights do not. As a result, the formulas below have $p = 1$ as a special case. (The $p = 1$ case where all the weights are equal to 1 is a straightforward limit of the $p > 1$ case; the weights $\alpha(e)$ become irrelevant in the limit $p \to 1$.)

DEFINITION A.2. For $1 < p \leqslant \infty$ and $f : G^p \to K$ a PL map from a p-conformal graph to a length graph, define

$$
E^p(f) := \begin{cases} \|f'\|_{p,G} & p > 1, \\ \int_G w(x) |f'(x)| \, dx & p = 1. \end{cases} \tag{A.3}
$$

Here, we take the L^p norm of $|f'|$ by integration with respect to α, and use an additional subscript to make clear where the norm is being evaluated. The metrics on G and K are used to evaluate the derivative $|f'|$, and the metric on G is also used in the integration in taking the L^p norm. Observe that the metric on G is irrelevant for $p = 1$.

If f is constant-derivative and $1 < p < \infty$, then this becomes

$$
E^p(f) = \left(\sum_{e \in \text{Edge}(G)} \frac{\ell(f(e))^p}{\alpha(e)^{p-1}} \right)^{1/p}.
$$

(Compare Equation (1.9).) For $f : W \to G^p$ a PL map from a weighted graph to a p-conformal graph, define

$$
E_p(f) := \begin{cases} \|n_f\|_{p^\vee, G} & p > 1, \\ \|n_f / w\|_{\infty, G} & p = 1, \end{cases} \tag{A.4}
$$

where $p^\vee = p/(p-1)$ is the Hölder conjugate of p. If n_f is constant on each

edge (as for taut maps, cf. Theorem 3) and $1 < p < \infty$, then

$$E_p(f) = \left(\sum_{e \in \mathrm{Edge}(G)} \alpha(e) n_f(e)^{p^\vee} \right)^{1/p^\vee}.$$

(Compare Equation (1.20).) In general, for $f : G^p \to H^q$ a PL map from a p-conformal graph to a q-conformal graph with $1 \leqslant p \leqslant q \leqslant \infty$ with $1 < q$ and $p < \infty$, define

$$\mathrm{Fill}^p(f) : H^q \to \mathbb{R}_{\geqslant 0}$$

$$\mathrm{Fill}^p(f)(y) := \begin{cases} \sum_{x \in f^{-1}(y)} |f'(x)|^{p-1} & p > 1, \\ \sum_{x \in f^{-1}(y)} w(x) & p = 1, \end{cases} \tag{A.5}$$

$$E_q^p(f) := \left(\|\mathrm{Fill}^p(f)\|_{q/(q-p),H} \right)^{1/p}. \tag{A.6}$$

If $p < q$, this is

$$E_q^p(f) = \left(\int_H \mathrm{Fill}^p(f)(y)^{1/(1-p/q)} \, d\alpha(y) \right)^{1/p - 1/q}. \tag{A.7}$$

Energies of homotopy classes are defined as an infimum as usual:

$$E_p[f] := \inf_{g \in [f]} E_p(g) \qquad E_q^p[f] := \inf_{g \in [f]} E_q^p(g) \qquad E^p[f] := \inf_{g \in [f]} E^p(g).$$

REMARK A.8. As with the earlier energies, in each case $E_q^p(f)$ naturally extends to a wider class of graph maps than PL maps, with the regularity required depending on p and q.

PROPOSITION A.9. *For a PL map f, we have $E_p(f) = E_p^1(f)$ for $1 < p \leqslant \infty$ and $E^p(f) = E_\infty^p(f)$ for $1 \leqslant p < \infty$.*

Proof. True by definition for E_p^1, and change of variables for E_∞^p. □

In light of Proposition A.9, we also define

$$E_\infty^\infty(f) := E^\infty(f) = \mathrm{Lip}(f)$$
$$E_1^1(f) := E_1(f) = \mathrm{WR}(f).$$

PROPOSITION A.10. *For $1 \leqslant p \leqslant q \leqslant \infty$ and $\phi : G^p \to H^q$ a PL map,*

$$E_q^p(\phi) = \sup_{f : H \to K} \frac{E^p(f \circ \phi)}{E^q(f)}$$

where the supremum runs over all metric graphs K and PL maps $f : H^q \to K$ with nonzero energy. If $p < q < \infty$, equality holds exactly when $|f'|$ is proportional to $\left(\mathrm{Fill}^p(\phi)\right)^{1/(q-p)}$.

Observe that Proposition A.10 is about energies of concrete maps, not about homotopy classes of maps.

Proof. We omit the cases when $p = 1$ and $q = \infty$. Otherwise, the energies $E^q(f)$ and $E^p(f \circ \phi)$ depend only on $|f'|$, so we may as well assume that K has the same underlying graph as H (with varying metric) and f is the identity as a graph map. We are thus picking $|f'|$ as a piecewise-constant function on H. We have

$$
\begin{aligned}
E^p(f \circ \phi) &= \left(\int_G \left| f'(\phi(x)) \right|^p |\phi'(x)|^p \, dx \right)^{1/p} \\
&= \left(\int_H \left| f'(y) \right|^p \left(\mathrm{Fill}^p(\phi)(y) \right) dy \right)^{1/p} \\
&\leqslant \| f' \|_q \left(\| \mathrm{Fill}^p(\phi) \|_{q/(q-p)} \right)^{1/p} \\
&= E^q(f) E^p_q(\phi),
\end{aligned}
$$

using Hölder's inequality in the form $\|ab\|_1 \leqslant \|a\|_{q/p} \|b\|_{q/(q-p)}$. The equality statement follows from the equality condition in Hölder's inequality. □

REMARK A.11. It is also true that

$$
E^p_q(\phi) = \sup_{c : W \to G} \frac{E_q(\phi \circ c)}{E_p(c)}
$$

where the supremum runs over all maps $c : W \to G$ from a weighted graph to G with nonzero energy (not necessarily taut). Here, equality holds when n_c is proportional to $|\phi'|^{p-1} \left(\mathrm{Fill}^p(\phi) \right)^{(p-1)/(q-p)}$.

PROPOSITION A.12. *Given $1 \leqslant p \leqslant q \leqslant r \leqslant \infty$ and a sequence $G_1^p \xrightarrow{f} G_2^q \xrightarrow{g} G_3^r$ of maps between a marked p-conformal graph G_1, a marked q-conformal graph G_2, and a marked r-conformal graph G_3,*

$$
\begin{aligned}
E^p_r(g \circ f) &\leqslant E^p_q(f) E^q_r(g) \\
E^p_r[g \circ f] &\leqslant E^p_q[f] E^q_r[g].
\end{aligned}
$$

Proof sketch. The first equation is an immediate consequence of Proposition A.10. The second equation follows as in the proof of Proposition 2.15. □

Taut maps automatically minimize E_p within their homotopy class, as in Proposition 3.8. We can also minimize E^p within a homotopy class.

PROPOSITION A.13. *Let* $[f]: G^p \to K$ *be a homotopy class of maps from a marked p-conformal graph to a marked length graph. Then there is a constant-derivative PL map* $g \in [f]$ *so that* $E^p(g) = E^p[f]$.

Proof sketch. For $p = \infty$, this is Theorem 4. For $p < \infty$, the form of E^p guarantees that an energy minimizer will be constant-derivative. As in Theorem 5, this reduces the space of possibilities to minimizing over a compact space for each of finitely many combinatorial types. □

A map that minimizes $E^p(f)$ within $[f]$ is called a *p-harmonic map*. Theorem 5′ may also be extended to give a local characterization of p-harmonic maps, including cases where the target is weak. Specifically, $[f]: G^p \to K$ is p-harmonic if and only if the map $W_f \to K$ is taut, where W_f is G with p-tension weights $w(x) = |f'(x)|^{p-1}$.

There are also versions of stretch factors: for $[\phi]: G^p \to H^q$ a homotopy class of marked maps, define

$$\overrightarrow{\mathrm{SF}}^p_q[\phi] := \sup_{[f]: H \to K} \frac{E^p[f \circ \phi]}{E^q[f]} \qquad (A.14)$$

$$\overleftarrow{\mathrm{SF}}^p_q[\phi] := \sup_{[c]: W \to G} \frac{E_q[\phi \circ c]}{E_p[c]}, \qquad (A.15)$$

where in (A.14) we maximize over marked length graphs K and PL maps $f: H^q \to K$ and in (A.15) we maximize over all marked weighted graphs (or multicurves) on G^p.

THEOREM 6. *For* $1 \leqslant p \leqslant q \leqslant \infty$ *and* $[\phi]: G^p \to H^q$ *a homotopy class of maps from a marked p-conformal graph to a marked q-conformal graph, there is a map* $\psi \in [\phi]$, *a marked weighted graph* W, *a marked weak length graph* K, *and a tight sequence of marked maps*

$$W \xrightarrow{\ c\ } G^p \xrightarrow{\ \psi\ } H^q \xrightarrow{\ f\ } K.$$

In particular,

$$E^p_q(\psi) = E^p_q[\phi] = \frac{E^p(f \circ \psi)}{E^q(f)} = \overrightarrow{\mathrm{SF}}^p_q[\phi] = \frac{E_q(\psi \circ c)}{E_p(c)} = \overleftarrow{\mathrm{SF}}^p_q[\phi].$$

Theorem 6 is much harder than Proposition A.10.

Proof sketch. The proof is quite similar to the proof of Theorem 1 in Section 6. For $p = 1$, the tautness results of Section 3 give the result, while for $q = \infty$ this is essentially Proposition A.13. So assume that $1 < p \leqslant q < \infty$.

For G^p a p-conformal graph with $1 < p < \infty$, there is an invertible 'duality' map $D_G^p \colon \mathcal{W}(G) \to \mathcal{L}(G)$, defined by setting

$$\big(D_G^p(w)\big)(e) = \alpha(e)w(e)^{1/(p-1)}. \tag{A.16}$$

Then, for a homotopy class as in the statement of Theorem 6, define an iteration $\mathrm{Iter}_\phi \colon \mathcal{L}(\Gamma_2) \to \mathcal{L}(\Gamma_2)$ as follows.

(1) For $\ell \in \mathcal{L}(\Gamma_2)$, set $K = (\Gamma_2, \ell)$, find a p-harmonic representative g of $[\mathrm{id} \circ \phi] \colon G^p \to K$, and set $m(e) = \ell(g(e)) \in \mathcal{L}(\Gamma_1)$.

(2) Set $w = (D_1^p)^{-1}(m) \in \mathcal{W}(\Gamma_1)$, so (Γ_1, w) is the tension-weighted graph of g.

(3) Set $v = N_{[\phi]}(w) \in \mathcal{W}(\Gamma_2)$.

(4) Set $\mathrm{Iter}_\phi(\ell) = D_2^q(v) \in \mathcal{L}(\Gamma_2)$.

If $p = q$, Iter_ϕ descends to a map on projective spaces $P\,\mathrm{Iter}_\phi \colon P\mathcal{L}(\Gamma_2) \to P\mathcal{L}(\Gamma_2)$, and one has to do an analysis of the possible boundary fixed points, entirely parallel to Section 6.3.

If $p < q$, then Iter_ϕ is not linear on rays, and we do the iteration on $\mathcal{L}(\Gamma_2)$ itself. More specifically, on a ray we have

$$\mathrm{Iter}_\phi(\lambda\ell) = \lambda^{(p-1)/(q-1)}\,\mathrm{Iter}_\phi(\ell).$$

Since $(p - 1)/(q - 1) < 1$, if a ray in $\mathcal{L}(\Gamma_2)$ is mapped to itself then there is a finite fixed point on the ray. A little more analysis shows that there are never attracting fixed points on the boundary of $\mathcal{L}(\Gamma_2)$, so there must be a fixed point in the interior. This gives strip graphs compatible with G^p and H^q in the sense of Definition A.17, along with a 1-filling map between them. $\qquad\square$

We now turn to the first alternate definition of p-conformal graphs and E_q^p. Recall from Definition 6.1 that a strip graph is tuple (Γ, w, ℓ) of a graph Γ and a set of weights w and lengths ℓ on Γ.

DEFINITION A.17. For $p \in [1, \infty)$, a *p-conformal rescaling* of a positive strip graph (Γ, w, ℓ) changes the weight, length, and area by

$$(w, \ell, \mathrm{Area}) \mapsto (\lambda^{p-1}w, \lambda\ell, \lambda^p\,\mathrm{Area})$$

where $\lambda \colon \text{Edge}(\Gamma) \to \mathbb{R}_+$ is a positive rescaling factor on each edge. (The identity $\text{Area} = w \cdot \ell$ is preserved.) An ∞-*conformal rescaling* instead acts by $(w, \ell, \text{Area}) \mapsto (\lambda w, \ell, \lambda \text{Area})$. We write $(\Gamma, w_1, \ell_1) \equiv_p (\Gamma, w_2, \ell_2)$ if the two strip structures are related by a p-conformal rescaling.

For $p \in (1, \infty]$, we say that a p-conformal graph (Γ, α) is *compatible* with a positive strip structure (w, ℓ) on Γ if

$$(\Gamma, w, \ell) \equiv_p (\Gamma, 1, \alpha),$$

or equivalently if $\ell(e) = \alpha(e) w(e)^{1/(p-1)}$. Thus we may think of a p-conformal graph as an equivalence class under $\equiv_p$. We say that (Γ, α) is compatible with an arbitrary (not necessarily positive) strip structure if, for each edge e,

$$\ell(e)^{p-1} = \alpha(e)^{p-1} w(e).$$

A 1-conformal graph is compatible with a strip structure if the weights agree (ignoring the lengths).

To better understand the duality map D^p from Equation (A.16), suppose we have a p-conformal graph (Γ, α). Then for $w \in \mathcal{W}^+(G)$, the lengths $D^p(w)$ are the unique values so that $(\Gamma, w, D^p(w)) \equiv_p (\Gamma, 1, \alpha)$.

DEFINITION A.18. Let $S_1 = (\Gamma_1, w_1, \ell_1)$ and $S_2 = (\Gamma_2, w_2, \ell_2)$ be two marked strip graphs (not necessarily balanced), with S_2 positive. For $\lambda > 0$, a map $f \colon S_1 \to S_2$ is *weakly λ-filling* if it satisfies Conditions (2) and (3) of Definition 6.4, dropping the condition that f is taut as a map between weighted graphs.

DEFINITION A.19. For $1 \leqslant p \leqslant q \leqslant \infty$, let G_1^p be a p-conformal graph, G_2^q be a q-conformal graph, and $f \colon G_1^p \to G_2^q$ be a PL map. Then $E_{q,\text{strip}}^p(f)$ is defined in the following way.

(1) If $p < q$, then, as we prove in Proposition A.22, there are strip graphs S_1 and S_2, compatible with subdivisions of G_1^p and G_2^q respectively, so that f is weakly 1-filling as a map from S_1 to S_2. Then

$$E_{q,\text{strip}}^p(f) = \text{Area}(S_2)^{1/p - 1/q}. \tag{A.20}$$

(2) If $p = q < \infty$, there are usually no strip structures that make f weakly 1-filling. Instead, take any strip structures (Γ_1, w_1, ℓ_1) and (Γ_2, w_2, ℓ_2) so that f is length-preserving, and define the energy to be the maximum ratio of weights:

$$E_{p,\text{strip}}^p(f) = \operatorname*{ess\,sup}_{y \in \Gamma_2} \frac{n_f^{w_1}(y)}{w_2(y)}. \tag{A.21}$$

This is independent of the choice of strip structures.

PROPOSITION A.22. *For $\phi\colon G_1^p \to G_2^q$ a PL map from a p-conformal graph to a q-conformal graph, $E_q^p(\phi) = E_{q,strip}^p(\phi)$.*

Proof sketch. This is closely related to Proposition A.10. For $p < q$, subdivide G_2 so that $\mathrm{Fill}^p(\phi)$ is constant on each edge. Then for e an edge of G_2, construct the strip structure (Γ_2, w_2, ℓ_2) compatible with G_2 so that

$$\ell_2(e) = \alpha_2(e) \cdot \left(\mathrm{Fill}^p(\phi)(e)\right)^{1/(q-p)}.$$

This determines ℓ_1 by the condition that ϕ be length-preserving, and w_1 and w_2 by the compatibility condition. It is elementary to check that ϕ is 1-filling with respect to these strip structures and then verify that $E_q^p(\phi) = E_{q,\mathrm{strip}}^p(\phi)$.

The case $p = q$ is easier, as you can choose ℓ_2 arbitrarily. $\qquad\square$

For the final variation on the definition of E_q^p, we allow more general spaces than graphs.

DEFINITION A.23. For $1 \leqslant p \leqslant \infty$, a *p-conformal space* is (loosely) a tuple (X, ℓ, μ) of a space X, a length metric ℓ on X, and a measure μ on X, up to rescaling by

$$(X, \ell, \mu) \equiv_p (X, \lambda\ell, \lambda^p\mu)$$

for a suitable rescaling function $\lambda\colon X \to \mathbb{R}_{>0}$. Write $[(X, \ell, \mu)]_p$ for an equivalence class of $\equiv_p$.

There are analytic subtleties in Definition A.23 in, for example, how to define the rescaling and exactly which metrics are allowed; we do not attempt to address them in this paper. But note that oriented conformal n-manifolds M^n give examples of n-conformal spaces: given a conformal class of (Riemannian) metrics on M, pick a base metric g in the conformal class, and set ℓ and μ to be distance with respect to g and the Lebesgue measure of g, respectively. Picking a different metric in the conformal class changes ℓ and μ by an n-conformal rescaling.

Suppose $X = \Gamma$ is a graph with a base metric and associated measure and:

- ℓ is a piecewise-constant multiple of the base metric;

- μ is a piecewise-constant multiple of the base Lebesgue measure; and

- the rescaling functions λ are piecewise-constant.

Definition A.23 is then almost identical to Definition A.17, if we define the weight at a generic point $x \in \Gamma$ by

$$w(x) = \frac{\mu(\Delta x)}{\ell(\Delta x)}$$

where Δx is a small interval centered on x.

DEFINITION A.24. For $1 \leqslant p \leqslant q \leqslant \infty$ with $p < \infty$ and $\phi \colon X_1^p \to X_2^q$ a suitable map from a p-conformal space to a q-conformal space, with $X_i = [(X_i, \ell_i, \mu_i)]$, define

$$\mathrm{Fill}_{\mathrm{conf}}^p(\phi) \colon Y^q \to \mathbb{R}_{\geqslant 0}$$
$$\mathrm{Fill}_{\mathrm{conf}}^p(\phi) := \phi_* \big(\big(\mathrm{Lip}_{\ell_2}^{\ell_1}(\phi)\big)^p \cdot \mu_1 \big) / \mu_2$$
$$E_{q,\mathrm{conf}}^p(\phi) := \big(\|\mathrm{Fill}_{\mathrm{conf}}^p(\phi)\|_{q/(q-p), X_2} \big)^{1/p}.$$

To take the definition of $E_{q,\mathrm{conf}}^p$ step by step:

- $\mathrm{Lip}_{\ell_2}^{\ell_1}(\phi) \colon X_1 \to \mathbb{R}_+$ is the local Lipschitz constant of ϕ.

- Next, $\phi_* \big(\big(\mathrm{Lip}_{\ell_2}^{\ell_1}(\phi)\big)^p \cdot \mu_1 \big)$ is the push-forward of measures.

- $\mathrm{Fill}_{\mathrm{conf}}^p(\phi) = \phi_*(\mathrm{Lip}(\phi)^p \cdot \mu_1)/\mu_2$ is the Radon–Nikodym derivative of the two measures.

- Finally, $E_{q,\mathrm{conf}}^p(\phi)$ is (up to a power) the $L^{q/(q-p)}$-norm of $\mathrm{Fill}_{\mathrm{conf}}^p$.

We will not define which maps ϕ are 'suitable' (or indeed which tuples (X, ℓ, μ) are allowed), but it should include cases where ϕ is Lipschitz and the Radon–Nikodym derivative exists, that is, $\phi_*(\mathrm{Lip}(\phi)^p \cdot \mu_1)$ is absolutely continuous with respect to μ_2. This includes nonconstant PL maps between graphs.

For $q = \infty$ (so that X_2 is a length space), $E_{\infty,\mathrm{conf}}^p$ can be rewritten as

$$E_{\infty,\mathrm{conf}}^p(\phi) = \big\| \mathrm{Lip}_{\ell_2}^{\ell_1}(\phi) \big\|_{p, X_1}. \tag{A.25}$$

In this case we do not need the Radon–Nikodym derivative.

The motivation for the exponents in Definition A.24 is that, up to an overall power, $E_{q,\mathrm{conf}}^p$ is the unique expression constructed with this data and these operations that is invariant under both p-conformal rescaling on X_1^p and q-conformal rescaling on X_2^q.

PROPOSITION A.26. *For $f\colon G^p \to H^q$ a PL map from a p-conformal graph to a q-conformal graph, $E_q^p(f) = E_{q,conf}^p(f)$.*

Proof. Expand the definitions. □

Definitions A.23 and A.24 point to a considerably more general setting, likely with substantial new difficulties. As mentioned in Section 1.4, much prior attention has been devoted to proving the existence of harmonic maps between various types of spaces [**EF01**] (related to minimizing E_∞^2), and the general case is likely to be harder.

WARNING A.27. The energy E_∞^2 from Definition A.24 does not agree with other definitions of Dirichlet energy. For instance, suppose X_1 is a Riemann surface Σ (with its natural 2-conformal structure), and X_2 is a Riemannian n-manifold M. Pick a base metric g on Σ in the given conformal class. Then, given a smooth map $f\colon \Sigma \to M$, we can consider the Jacobian $df_x\colon T_x\Sigma \to TM$, with singular values $\lambda_1, \lambda_2\colon \Sigma \to \mathbb{R}_{\geqslant 0}$, the eigenvalues of $\sqrt{(df_x)^T(df_x)}$, chosen so that $\lambda_1(x) \geqslant \lambda_2(x)$. Thus df_x maps the unit circle in $T_x\Sigma$ to an ellipse whose major and minor axes have length $\lambda_1(x)$ and $\lambda_2(x)$. The local Lipschitz constant of f at x is $\lambda_1(x)$, so we have

$$\left(E_{\infty,\mathrm{conf}}^2(f)\right)^2 = \int_\Sigma \lambda_1(x)^2\,\mu(x)$$

(where μ is the Lebesgue measure on Σ) while the standard Dirichlet energy is

$$\mathrm{Dir}(f) = \int_\Sigma \left(\lambda_1(x)^2 + \lambda_2(x)^2\right)\mu(x).$$

These energies are both conformally invariant, but are not the same. They do agree if the target space is a graph, as in that case $\lambda_2(x) = 0$.

Appendix B. Electrical networks

As mentioned in the introduction, the elastic graphs of this paper are closely related to the much better studied theory of *resistor networks*. Suppose we are given an elastic graph G with k marked vertices $x_1, \ldots, x_k$, called *nodes*. Turn it into a network of resistors, where the elastic constants $\alpha(e)$ become resistances. If we attach external voltage sources at voltages $V_1, \ldots, V_k$ to the nodes, then the remainder of the circuit will reach an electrical equilibrium, which has several pieces of data:

- a voltage $V(v)$ for each vertex v of G (agreeing with $V_1, \ldots, V_k$ on the nodes);

- an internal current $I(\vec{e})$ flowing through each oriented edge $\vec{e}$ of G (with $I(-\vec{e}) = -I(\vec{e})$);

- the total current $I_1, \ldots, I_k$ flowing out of the nodes; and

- the total energy E dissipated by the system per unit time.

At equilibrium, these are related by Kirchhoff's current laws.

- The current on an edge is related to the voltage difference. If $\vec{e}$ has source s and target t, let $\Delta V(\vec{e}) = V(t) - V(s)$; then

$$I(\vec{e}) = \frac{\Delta V(\vec{e})}{\alpha(e)}.$$

- For each internal (unmarked) vertex v of G, the total current flowing in is 0

$$\sum_{\vec{e} \text{ incident to } v} I(\vec{e}) = 0,$$

while at the node x_i,

$$I_i = \sum_{\vec{e} \text{ incident to } x_i} I(\vec{e}).$$

- The energy dissipated is

$$E = \sum_{e \in \text{Edge}(G)} \alpha(e) I(\vec{e})^2 = \sum_{e \in \text{Edge}(G)} \frac{(\Delta V(\vec{e}))^2}{\alpha(e)}.$$

The energy dissipated is identical to Equation (1.9) for the Dirichlet energy of a map f, in the special case that the target of f is $\mathbb{R}$ with k marked points at $V_1, \ldots, V_k$.

For resistor networks, the equations for the internal voltages and currents are linear, so $V(v)$, $I(\vec{e})$, and I_i are linear functions of the V_i, while E is a quadratic function of the V_i. (By contrast, in the more general case considered in the bulk of this paper, the energy $\text{Dir}_{[f]}$ as a function of lengths is only piecewise-quadratic.) The *response matrix* $\Lambda_{ij}(G)$ of a resistor network G is the matrix that gives the external currents I_i as a function of the external voltages V_j. The matrix Λ is symmetric and determines E as a quadratic function of the V_i:

$$E = \sum_{i,j=1}^{k} V_i V_j \Lambda_{ij}. \tag{B.1}$$

Much attention has been devoted to the question of when two resistor networks (with the same number of nodes) are *electrically equivalent*, in the sense that the response matrices are the same. Series and parallel reductions of resistors are examples of electrical equivalence. A more substantial example [**Ken99**] is the $Y-\Delta$ *transform*, that relates a 3-node network with the topology of a Y to one with the topology of a Δ:

$$\equiv_{\text{elec}} \tag{B.2}$$
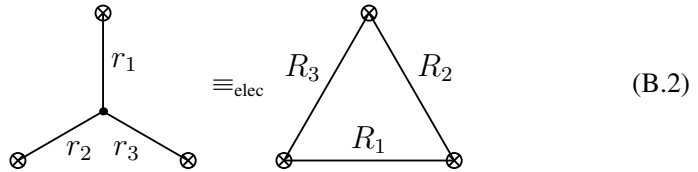
where

$$r_1 = \frac{R_2 R_3}{R_1 + R_2 + R_3} \qquad\qquad R_1 = \frac{r_2 r_3 + r_1 r_3 + r_1 r_2}{r_1}$$

$$r_2 = \frac{R_1 R_3}{R_1 + R_2 + R_3} \qquad\qquad R_2 = \frac{r_2 r_3 + r_1 r_3 + r_1 r_2}{r_2}$$

$$r_3 = \frac{R_1 R_2}{R_1 + R_2 + R_3} \qquad\qquad R_3 = \frac{r_2 r_3 + r_1 r_3 + r_1 r_2}{r_3}$$

It turns out that the $Y-\Delta$ transform and series and parallel reductions are sufficient to relate any two electrically equivalent planar resistor networks with nodes on the external face [**CdV94**, **CdVGV96**, **CIM98**].

We can also ask when one resistor network G_1 *dominates* another network G_2, in the sense that the energy dissipated by G_1 is greater than the energy dissipated by G_2, for any choice of external voltages V_i. From Equation (B.1) we see that this happens exactly when $\Lambda(G_1) \preceq \Lambda(G_2)$, in the usual Löwner ordering on quadratic forms.

Elastic networks with targets more general than $\mathbb{R}$ are almost never equivalent, so instead we ask about domination, as in Theorem 1. If G_1, G_2 are elastic networks with k corresponding marked nodes, say that $G_1 \preceq_{\text{elast}} G_2$ if, for any maps ϕ_i from G_i to a marked tree K mapping corresponding nodes to the same point,

$$\text{Dir}[\phi_1] \leqslant \text{Dir}[\phi_2].$$

Then, for instance, we have the following inequalities of energies:

$$\leq_{\text{elast}} \qquad \leq_{\text{elast}} \qquad , \qquad \text{(B.3)}$$

The second two graphs are electrically equivalent by (B.2). The first inequality is a simple application of Theorem 1, while the second requires a little more argument. Since the second and third graphs are electrically equivalent, and those two graphs have equal energies when K is $\mathbb{R}^n$ (with three marked points) for any n. (The case $n = 1$ is exactly electrical equivalence, and for $n \geqslant 2$ the Dirichlet energy is the sum of the energies of the projections to the different coordinates.) In fact, the inequalities in (B.3) hold more generally when K is any CAT(0) space with three marked points.

More generally, a connected resistor network G with three nodes is always electrically equivalent to a tripod with some weights and a triangle with some other weights, related to each other by the $Y-\Delta$ transform. The tripod is always elastically dominated by the electrically equivalent triangle. In general, in forthcoming joint work with Dejean and Gorski, we have the following theorem.

THEOREM B.4 (Dejean–Gorski–Thurston [**DGT**]). *Let G be an elastic network with three nodes, and let Y_G and Δ_G be the electrically equivalent tripod and triangle, respectively. Then*

$$Y_G \leq_{elast} G \leq_{elast} \Delta_G.$$

We close by reminding the reader of the connection between electrical networks at equilibrium and rectangle tilings [**BSST40**]: loosely, if you assign each edge a rectangle of length equal to the voltage difference between the endpoints and width equal to the total current, then Kirchhoff's laws say that the rectangles may be assembled into a single tiling, in which the aspect ratios are equal to the resistances. See Figure 6 for a simple example. In the more general setting of elastic graphs, the 'weights' throughout this paper can also be reinterpreted as 'widths', giving similar tiling pictures on a surface.

Let us spell this out a bit more, although without proofs, and omitting background on quadratic differentials. Suppose that $f \colon G \to K$ is a surjective harmonic map from an elastic graph to a length graph, and suppose that G and K both have a ribbon structure and f is a ribbon map, that is, extends to an embedding $Nf \colon NG \hookrightarrow NK$ between surface thickenings of G and K [**Thu16b**, Definition 2.7]. Let $\Sigma = NK$ be the thickening of K. Recall that a
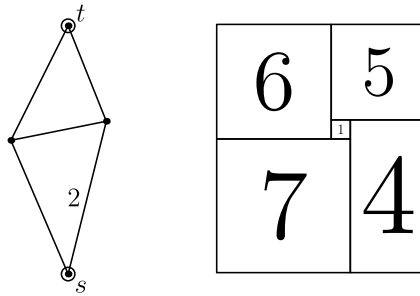
Figure 6. A simple electric network and an associated tiling by rectangles. All resistances on the graph are 1 except for one edge, which has resistance 2. On the rectangle tiling, we have shown the total current flowing through the edge, which in the picture is the width of the rectangles.

flip-translation structure on Σ is an atlas with overlap maps given by translations or rotations by π, with certain allowed singularities; equivalently, it is a pair of a conformal structure ω on Σ and a quadratic differential q with respect to ω. Then there is a flip-translation structure on Σ with the following properties.

(1) The boundaries of Σ are horizontal.

(2) Corresponding to each vertex of K of valence v, there is a zero of q of order $v - 2$. (Note that K cannot have vertices of valence 1.)

(3) The quotient of q by the vertical foliation is the length graph K, with length induced by the measure on the vertical foliation.

(4) Σ is tiled by rectangles (axis-aligned in the flip-translation structure), with each edge e corresponding to a rectangle $R(e)$. The aspect ratio of $R(e)$ is $\alpha(e)$, while the length (horizontal extent) of $R(e)$ is equal to the length of $f(e)$ in K.

(5) The measure on the horizontal foliation of q gives weights related to a weighted graph W_f forming a tight sequence

$$W \longrightarrow G \xrightarrow{f} K$$

in the sense of Theorem 5. It is tempting to think that the leaves of the horizontal foliation itself give the weighted multicurve C in Theorem 5, but this is not generally true, as the leaves of the horizontal foliation will typically not be closed when K is a sufficiently complicated graph.

The fact that a harmonic map gives a quadratic (not holomorphic) differential on Σ, or equivalently the fact that we have a flip-translation structure rather

than a translation structure on Σ, comes from the distinction between electrical and elastic stretching mentioned in the introduction: an electrical flow through a network gives an orientation on each edge, while a stretched elastic graph has unsigned tensions on each edge.

References

[BS17]   M. C. Bell and S. Schleimer, 'Slow north-south dynamics on $\mathcal{PML}$', *Groups Geom. Dyn.* **11**(3) (2017), 1103–1112.

[Bes11]   M. Bestvina, 'A Bers-like proof of the existence of train tracks for free group automorphisms', *Fund. Math.* **214**(1) (2011), 1–12.

[BH92]   M. Bestvina and M. Handel, 'Train tracks and automorphisms of free groups', *Ann. of Math. (2)* **135**(1) (1992), 1–51.

[BSST40]   R. L. Brooks, C. A. B. Smith, A. H. Stone and W. T. Tutte, 'The dissection of rectangles into squares', *Duke Math. J.* **7** (1940), 312–340.

[CFP94]   J. W. Cannon, W. J. Floyd and W. R. Parry, 'Squaring rectangles: the finite Riemann mapping theorem', in *The Mathematical Legacy of Wilhelm Magnus: Groups, Geometry and Special Functions (Brooklyn, NY, 1992)*, Contemporary Mathematics, 169 (American Mathematical Society, Providence, RI, 1994), 133–212.

[CS11]   D. Chelkak and S. Smirnov, 'Discrete complex analysis on isoradial graphs', *Adv. Math.* **228**(3) (2011), 1590–1630.

[CdV94]   Y. C. de Verdière, 'Réseaux électriques planaires. I', *Comment. Math. Helv.* **69**(3) (1994), 351–374.

[CdVGV96]   Y. C. de Verdière, I. Gitler and D. Vertigan, 'Réseaux électriques planaires. II', *Comment. Math. Helv.* **71**(1) (1996), 144–167.

[CIM98]   E. B. Curtis, D. Ingerman and J. A. Morrow, 'Circular planar graphs and resistor networks', *Linear Algebra Appl.* **283**(1–3) (1998), 115–150.

[DGT]   B. Dejean, C. Gorski and D. Thurston, 'Dirichlet energies of 3-marked elastic graphs'. in preparation, Based on an REU project, http://www.math.indiana.edu/reu/2016/reu 2016.pdf.

[DH93]   A. Douady and J. H. Hubbard, 'A proof of Thurston's topological characterization of rational functions', *Acta Math.* **171**(2) (1993), 263–297.

[Duf62]   R. J. Duffin, 'The extremal length of a network', *J. Math. Anal. Appl.* **5** (1962), 200–215.

[Duf68]   R. J. Duffin, 'Potential theory on a rhombic lattice', *J. Combin. Theory* **5** (1968), 258–272.

[EF01]   J. Eells and B. Fuglede, *Harmonic Maps between Riemannian Polyhedra*, Cambridge Tracts in Mathematics, 142 (Cambridge University Press, Cambridge, 2001).

[Fer44]   J. Ferrand, 'Fonctions préharmoniques et fonctions préholomorphes', *Bull. Sci. Math. (2)* **68** (1944), 152–180.

[FM11]   S. Francaviglia and A. Martino, 'Metric properties of outer space', *Publ. Mat.* **55**(2) (2011), 433–473.

[Fug57]   B. Fuglede, 'Extremal length and functional completion', *Acta Math.* **98** (1957), 171–219.

[GS92]   M. Gromov and R. Schoen, 'Harmonic maps into singular spaces and p-adic superrigidity for lattices in groups of rank one', *Publ. Math. Inst. Hautes Études Sci.* **76** (1992), 165–246.

[HS96]   Z.-X. He and O. Schramm, 'On the convergence of circle packings to the Riemann map', *Invent. Math.* **125**(2) (1996), 285–305.

[Isa41]   R. P. Isaacs, 'A finite difference function theory', *Univ. Nac. Tucumán. Revista A.* **2** (1941), 177–201.

[Kah06]   J. Kahn, '*A priori* bounds for some infinitely renormalizable quadratics: I. Bounded primitive combinatorics', Preprint ims06-05, Stony Brook IMS, 2006, arXiv:math/0 609045v2.

[KPT15]   J. Kahn, K. M. Pilgrim and D. P. Thurston, 'Conformal surface embeddings and extremal length', Preprint, 2015, arXiv:1507.05294.

[Ken99]   A. E. Kennelly, 'Equivalence of triangles and stars in conducting networks', *Electrical World and Engineer* **34** (1899), 413–414.

[Ker80]   S. P. Kerckhoff, 'The asymptotic geometry of Teichmüller space', *Topology* **19**(1) (1980), 23–41.

[KMT03]   S. Kojima, S. Mizushima and S. P. Tan, 'Circle packings on surfaces with projective structures', *J. Differential Geom.* **63**(3) (2003), 349–397.

[KS93]   N. J. Korevaar and R. M. Schoen, 'Sobolev spaces and harmonic maps for metric space targets', *Comm. Anal. Geom.* **1**(3–4) (1993), 561–659.

[Lov04]   L. Lovász, 'Discrete analytic functions: an exposition', in *Surveys in Differential Geometry,* Vol. IX, Surveys in Differential Geometry, 9 (Int. Press, Somerville, MA, 2004), 241–273.

[Mer01]   C. Mercat, 'Discrete Riemann surfaces and the Ising model', *Comm. Math. Phys.* **218**(1) (2001), 177–216.

[Pal15]   D. R. Palmer, 'Toward computing extremal quasiconformal maps via discrete harmonic measured foliations', A.B. thesis, Harvard University, Cambridge, MA, November 2015.

[PH92]   R. C. Penner and J. L. Harer, *Combinatorics of Train Tracks*, Annals of Mathematics Studies, 125 (Princeton University Press, Princeton, NJ, 1992).

[Pic05]   J. Picard, 'Stochastic calculus and martingales on trees', *Ann. Inst. Henri Poincaré Probab. Stat.* **41**(4) (2005), 631–683.

[PP93]   U. Pinkall and K. Polthier, 'Computing discrete minimal surfaces and their conjugates', *Exp. Math.* **2**(1) (1993), 15–36.

[RS87]   B. Rodin and D. Sullivan, 'The convergence of circle packings to the Riemann mapping', *J. Differential Geom.* **26**(2) (1987), 349–360.

[Smi01]   S. Smirnov, 'Critical percolation in the plane', Preprint, 2001, arXiv:0909.4499.

[Smi10]   S. Smirnov, 'Discrete complex analysis and probability', in *Proceedings of the International Congress of Mathematicians,* Vol. I (Hindustan Book Agency, New Delhi, 2010), 595–621.

[Ste05]   K. Stephenson, *Introduction to Circle Packing* (Cambridge University Press, Cambridge, 2005).

[Thu16a]   D. P. Thurston, 'From rubber bands to rational maps: a research report', *Res. Math. Sci.* **3** (2016), Art. 15, arXiv:1502.02561.

[Thu16b]   D. P. Thurston, 'A positive characterization of rational maps', Preprint, 2016, arXiv: 1612.04424.

[Thu86]   W. P. Thurston, 'Zippers and univalent functions', in *The Bieberbach Conjecture (West Lafayette, Ind., 1985)*, Mathematical Surveys and Monographs, 21 (American Mathematical Society, Providence, RI, 1986), 185–197.

[Wol95]   M. Wolf, 'On the existence of Jenkins–Strebel differentials using harmonic maps from surfaces to graphs', *Ann. Acad. Sci. Fenn. Ser. A I Math.* **20**(2) (1995), 269–278.