FROM RUBBER BANDS TO RATIONAL MAPS
RESEARCH REPORT

DYLAN P. THURSTON

Dedicated to the memory of William P. Thurston, 1946–2012

Abstract. We develop parallel theories of elastic graphs and conformal surfaces with boundary. One one hand, this lets us tell when one rubber band network is looser than another, and on the other hand tell when one conformal surface embeds in another.

We apply this to give a new characterization of hyperbolic critically finite rational maps among branched self-coverings of the sphere, by a positive criterion: a branched covering is equivalent to a hyperbolic rational map if and only if there is an elastic graph with a particular “self-embedding” property. This complements the earlier negative criterion of W. Thurston.

 Portions of this project are joint work with Jeremy Kahn and Kevin Pilgrim.

Contents

1. Introduction 2
  1.1. Detecting rational maps 3
  1.2. Elastic graphs 6
  1.3. History and prior work 9
  1.4. Organization 9
  1.5. Acknowledgements 9

2. Examples 10
  2.1. Polynomials: The rabbit and the basilica 10
  2.2. Matings 11
  2.3. Slit maps 12
  2.4. Behavior under iteration 13

3. Setting 14
  3.1. Surfaces 14
  3.2. Convexity 15
  3.3. Graphs 17
  3.4. Extensions 19

4. Harmonic maps and Dirichlet energy 19
  4.1. Harmonic maps from surfaces 20
  4.2. Harmonic maps from graphs 20
  4.3. Relation to Lipshitz energy 22
  4.4. Computing Dirichlet energy 23

5. Extremal length 23
  5.1. Extremal length on surfaces 23
  5.2. Extremal length on graphs 24
1. Introduction

This paper is devoted to explaining a circle of ideas, relating:

- elastic networks (“rubber bands”) and the corresponding Dirichlet energy,
- extremal length and other “quadratic” norms on the space of curves in a surface,
- embeddings between Riemann surfaces, conformal or quasi-conformal, and
- post-critically finite rational maps.

The original motivation for this project is the last point, and more specifically the question of when a topological branched self-covering of the sphere is equivalent to a rational map. William Thurston first answered this question more than 30 years ago [DH93], by giving a “negative” characterization: a combinatorial obstruction that exists exactly when the maps is not rational (Theorem 8.16). In this paper, we give a “positive” characterization: a combinatorial object that exists exactly when the map is rational, for a somewhat restricted class of maps (Theorem 1).

Having both positive and negative combinatorial characterizations for the same property automatically gives an algorithm for testing the property. You may search, in parallel, for either an obstruction or a certificate for the property. One of the searches will eventually succeed and answer the question. There are other algorithms for
testing whether a map is rational: you may take the rational map itself as a certificate for being rational [BBY12]. However, we expect our combinatorial certificate to be practical to search for (Section 6.6). By contrast, W. Thurston’s obstruction theorem is notoriously hard to apply.

In addition, a positive characterization gives an object to study associated to the topological branched covers of most interest, namely the rational ones.

Several of the constructions along the way are of independent interest. For instance, we give a new characterization of when one Riemann surface conformally embeds in another in a given homotopy class, and a numerical invariant of such embeddings.

This is a preliminary report on the results. Most proofs are omitted, although we give some indications. We also indicate some of the many open problems suggested by this research in “Extensions” subsections; all of these are unnecessary for the main results.

1.1. Detecting rational maps. Recall that a (topological) branched self-cover of the sphere is a finite set of points $P$ in the sphere and a map $f: (S^2, P) \to (S^2, P)$ that is a covering when restricted to $S^2 \setminus f^{-1}(P)$. One central question is when such a map is equivalent to a rational map on $\mathbb{CP}^1$. (See Definition 8.1.) A branched self-cover can be characterized (up to homotopy) by picking a spine $\Gamma$ for $S^2 \setminus P$, and drawing its inverse image $\tilde{\Gamma} = f^{-1}(\Gamma) \subset S^2 \setminus f^{-1}(P)$. There are two natural homotopy classes of maps from $\tilde{\Gamma}$ to $\Gamma$.

- A covering map commuting with the action of $f$:

$$
\begin{array}{ccc}
\tilde{\Gamma} & \xrightarrow{f} & \Gamma \\
| & | & | \\
S^2 \setminus f^{-1}(P) & \xrightarrow{\pi} & S^2 \setminus P.
\end{array}
$$

We denote this covering map $\pi: \tilde{\Gamma} \to \Gamma$.

- A map commuting up to homotopy with the inclusion in $S^2 \setminus P$:

$$
\begin{array}{ccc}
\tilde{\Gamma} & \xrightarrow{\phi} & \Gamma \\
| & | & | \\
S^2 \setminus f^{-1}(P) & \xleftarrow{\text{id}} & S^2 \setminus P.
\end{array}
$$

We denote this map $\phi$. It is unique up to homotopy since $\Gamma$ is a spine for $S^2 \setminus P$.

Example 1.1. Figure 1 shows a simple example of such spines when $P$ has 3 points, $A$, $B$, and $C$ (at $\infty$). The covering map $\pi$ is preserves the colors and arrows. The branched self-cover $f: (S^2, P) \to (S^2, P)$ is an extension of $\pi$ to the whole sphere. This extension is unique up to homotopy relative to $\{A, B, C\}$. The map $f$ permutes the
Figure 1. A spine and its inverse image defining a branched self-cover of the sphere. The marked set \( P \) consists of the two points shown, and a third point \( C \) at infinity. The numbers on the edges give a measure which proves this branched self-cover is equivalent to a rational map.

marked points by

\[
\begin{array}{c}
A \\
\searrow
\end{array} \quad (2) \quad \begin{array}{c}
\swarrow
B
\end{array} \\
\begin{array}{c}
(2)
C
\end{array}
\]

(For instance, \( A \) is inside a red-green region in \( \tilde{\Gamma} \), so must map to \( B \), which is inside a red-green region in \( \Gamma \).)

The map \( \phi \), on the other hand, is the projection of \( \tilde{\Gamma} \) onto \( \Gamma \) considered as a spine; for instance, it might map the right-hand blue edge of \( \tilde{\Gamma} \) to the right-hand green edge of \( \Gamma \), whereas \( \pi \) preserves the colors.

For our characterization of rational maps, we also need a elastic structure on \( \Gamma \), by which we mean a measure \( \alpha \) on \( \Gamma \) absolutely continuous with respect to Lebesgue measure.\(^1\) We can pull back \( \alpha \) by \( \pi \) to get a measure \( \tilde{\alpha} \) on \( \tilde{\Gamma} \). For \( \psi : \tilde{\Gamma} \rightarrow \Gamma \) a Lipshitz map of graphs, define the embedding energy by

\[
\text{Emb}(\psi) := \text{ess sup}_{y \in \Gamma} \sum_{x \in \tilde{\Gamma}} \left| \psi'(x) \right|.
\]

The derivative is taken with respect to the two measures \( \tilde{\alpha} \) and \( \alpha \). The essential supremum \( \text{ess sup} \) ignores points of measure zero. In particular, ignore vertices of \( \tilde{\Gamma} \) or points of \( \tilde{\Gamma} \) that map to vertices of \( \Gamma \), where the derivative requires some thought to define.

In practice the embedding energy is optimized when \( \psi \) is piecewise-linear, and the reader may restrict to that case.

For one motivation for Equation (1.2), see Figure 6 and Proposition 6.38: the embedding energy characterizes conformal embeddings of thickened versions of the graph. Another one is in Theorem 6: it characterizes when Dirichlet energy is reduced under the map, i.e., when one rubber band network is looser than another. (See Remark 6.30.)

Example 1.3. To return to Example 1.1, consider the measure \( \alpha \) on \( \Gamma \) from Figure 1 and the concrete map \( \psi : \tilde{\Gamma} \rightarrow \Gamma \) that maps

- the right blue edge of \( \tilde{\Gamma} \) to the green edge on the right of \( \Gamma \),

\(^1\)We can also interpret \( \alpha \) as a metric. But we prefer to distinguish this measure from other metrics that come later.
• the right green edge of \( \Gamma \) to the red edge in the middle of \( \Gamma \), and
• the remaining four edges (two red, one green, and one blue) of \( \Gamma \) map to the blue edge on the right of \( \Gamma \), with the red edges mapping to segments of length \( 1/\sqrt{2} \).

Make \( \psi \) linear on each segment described above. Then a short computation shows that \( \text{Emb}(\psi) = 1/\sqrt{2} \).

A major theorem of this paper is that rational maps can be characterized by maps \( \psi \) with embedding energy less than 1. We say that two branched self-covers \( f_1: (S^2, P_1) \bowtie \) and \( f_2: (S^2, P_2) \bowtie \) are equivalent if they can be connected by a homotopy of \( f_1 \) relative to \( P_1 \) and conjugacy taking \( P_1 \) to \( P_2 \).

**Theorem 1.** Let \( f: (S^2, P) \bowtie \) be a branched self-cover of the sphere. Suppose that there is a branch point in each cycle in \( P \). Then \( f \) is equivalent to a rational map iff there is an elastic graph spine \( (\Gamma, \alpha) \) for \( S^2 \setminus P \), an integer \( n > 0 \), and map \( \psi \in [\phi_n] \) so that \( \text{Emb}(\psi) < 1 \).

Here \( \phi_n \) is obtained from \( \phi \) by iteration, and is in the homotopy class of the projection \( f^{-n}(\Gamma) \to \Gamma \). Loosely speaking, Theorem 1 says that the self-cover is rational iff there is a self-embedded spine for \( S^2 \setminus P \).

**Remark 1.4.** It is likely the condition on Theorem 1 can be relaxed to assume merely that \( f \) has at least one branch point in one cycle in \( P \), i.e., if \( f \) is rational, its Julia set is not the whole sphere. See Section 8.7.

**Example 1.5.** The explicit measure and map in Example 1.3 shows that the given branched self-cover is equivalent to a rational map. In fact, it is equivalent to the rational map

\[
f(z) = \frac{1}{1 - z^2},
\]

with \( A = 0 \), \( B = 1 \), and \( C = \infty \).\(^2\)

There was a previous characterization of rational maps by W. Thurston, recalled as Theorem 8.16 below. This is analogous to the two ways to characterize pseudo-Anosov surface automorphisms, which form a natural class of geometric elements of the mapping class group of a surface. Geometrically, pseudo-Anosov diffeomorphisms are those whose mapping torus is a hyperbolic 3-manifold. Combinatorially, there are two criteria:

• **Negative:** \( f \) is pseudo-Anosov iff it is not periodic (\( f^{\circ k} \) is the identity for some \( k > 0 \)) or reducible (there is an invariant system of multi-curves for \( f \)).
• **Positive:** \( f \) is pseudo-Anosov iff there is a measured train track \( T \) and a splitting sequence from \( T \) to a train-track \( T' \) with \( T' = \lambda f(T) \) for some constant \( 0 < \lambda < 1 \) [PP87].

The positive criterion gives some extra information: the number \( \lambda \) is an invariant of \( f \), with dynamical interpretations. (For instance, \( \lambda \) controls the growth rate of intersection numbers.)

\(^2\)Every branched self-cover with only 3 post-critical points is equivalent to a rational map, so this example was easy to do by other means.
Analogously, we can say that a branched self-cover $f$ of $(S^2, P)$ is geometric if it is equivalent to a rational map, which also have associated 3-dimensional hyperbolic laminations [LM97]. There are combinatorial criteria for $f$ to be rational:

- **Negative**: $f$ is rational iff there is no obstruction, as in W. Thurston’s Theorem 8.16. In loose terms, the obstruction is a back-expanding annular system: a collection of annuli that get “wider” under backwards iteration.

- **Positive**: Under some additional assumptions, $f$ is rational iff there is a metric spine for $S^2 \setminus P$ satisfying a back-contracting condition. This is Theorem 1.

As in the case of surface automorphisms, the two theorems are in a sense “dual” to each other: it is easy to see that a branched-self cover cannot simultaneously have a back-expanding annular system and a back-contracting spine. (See Equation (8.18).) Also as in the surface automorphism case, the positive criterion gives us a new object to study, namely the constant $\overline{\text{SF}}[\phi]$ of Section 7.2.

Compared to the situation for surface automorphisms, Theorem 1 has the following caveats:

- It only works when there is a branch point in each cycle in $P$. (But see Remark 1.4.)

- We may need to pass to an iterate to get $\text{Emb}(\phi) < 1$. Furthermore, it is easy to see in examples that the embedding energy can decrease in powers. (In the notation of Section 7.2, in general $\overline{\text{SF}}[\phi_\alpha n] < \overline{\text{SF}}[\phi]^n$ and $\overline{\text{SF}}[\phi] < \overline{\text{SF}}[\phi]$.)

There is a further caveat for both surface automorphisms and branched self-covers:

- For the positive criterion, the train track or graph constructed is not canonical: there are many different choices that work for the criterion.

By contrast, the negative criteria can be made canonical. (Pilgrim [Pil01] proved this for branched self-covers.) On the other hand, in the surface case, Agol [Ago11] and Hamenstädt [Ham09] give a canonical object related to the measured train track.

1.2. **Elastic graphs.** The embedding energy of Equation (1.2) looks a little mysterious; it looks a little like the Lipshitz stretch factor, but the sum over inverse images looks unusual. To explain where it comes from, we now turn to a “conformal” theory of graphs parallel to the conformal theory of Riemann surfaces. The central object is an elastic graph $(\Gamma, \alpha)$, which you should think of as a network of rubber bands; formally, it is a graph with a measure on each edge, representing the elasticity of the edge. (See Section 3.3.)

There are additional structures we can put on the graph.

- On one hand, we can consider curves $C$ on the graph, maps from a 1-manifold into $\Gamma$.

- On the other hand, we can consider maps from $\Gamma$ to a length graph $K$, a graph with fixed lengths of edges (like a network of pipes).
Maps between these objects have naturally associated energies, as summarized in the following diagram.

![Diagram](image)

The labels on the arrows indicate the type of energy on a map of this type, as follows.

- For a map \( f \) from an elastic graph \((\Gamma, \alpha)\) to a length graph \((K, \ell)\), there is the Dirichlet or rubber-band energy (Section 4) familiar from physics:

\[
\text{Dir}(f) := \int_{x \in \Gamma} |f'(x)|^2 \, dx,
\]

where \( f' \) measure the derivative with respect to the natural metrics. If \( f \) minimizes this energy within its homotopy class, it is said to be harmonic.

- For a curve \( C \) in an elastic graph \((\Gamma, \alpha)\), we have a version of extremal length (Section 5):

\[
\text{EL}[C] := \sum_{e \in \text{Edge}(\Gamma)} n_C(e)^2 \cdot \alpha(e),
\]

where \( n_C(e) \) is the number of times \( C \) runs over the edge \( e \) (without backtracking).

- For a curve \( C \) in a length graph \((K, \ell)\), we have the usual length, which in our notation is

\[
\ell[C] := \sum_{e \in \text{Edge}(K)} n_C(e) \cdot \ell(e).
\]

To match the other quantities, we actually use the square of the length as our energy.

- For a map \( \phi \) between length graphs \((K_1, \ell_1)\) and \((K_2, \ell_2)\), there is the Lipshitz constant (Section 4.3):

\[
\text{Lip}(\phi) := \text{ess sup}_{x \in K_1} |\phi'(x)|.
\]

Again, we consider the square of the Lipshitz energy.

- Finally, for a map \( \phi \) between elastic graphs \((\Gamma_1, \alpha_1)\) and \((\Gamma_2, \alpha_2)\), we have embedding energy as used in Theorem 1 (Section 6):

\[
\text{Emb}(\phi) := \text{ess sup}_{y \in \Gamma_2} \sum_{x \in \phi^{-1}(y)} |\phi'(x)|.
\]

This energy appears to be new, although it is related to Jeremy Kahn’s notion of domination of weighted arc diagrams (Section 8.5).

Remark 1.12. We could make the diagram more symmetric by using width graphs instead of curves (see Section 3.3), and adding a norm on maps between width graphs.
These energies are sub-multiplicative, in the sense that composing two maps can only decrease the product of the energies: if $f$ and $g$ are two composable maps of the above types, then

$$\|f \circ g\| \leq \|f\| \cdot \|g\|,$$

(1.13)

where $\|\cdot\|$ is the appropriate energy from the above list. (This inequality is the reason we squared some of the energies.) For instance, if we have fix elastic graphs $\Gamma_1$ to $\Gamma_2$, a length graph $K$, and maps $\phi: \Gamma_1 \to \Gamma_2$ and $f: \Gamma_2 \to K$, then

$$\text{Dir}(f \circ \phi) \leq \text{Emb}(f) \cdot \text{Dir}(\phi),$$

(1.14)

What is more, these inequalities are all tight, in the sense that if we fix the domain, range, and homotopy type of $f$, then we can find a sequence of functions $g_i$ (including a choice of domain) that approach equality in Equation (1.13). Likewise if we fix $g$ and vary $f$.

For instance, if $\phi: \Gamma_1 \to \Gamma_2$ is a map of elastic graphs, we can strengthen Equation (1.14) to

$$\text{Emb}[\phi] = \sup_{K \text{ length graph}} \frac{\text{Dir}[f \circ \phi]}{\text{Dir}[f]},$$

(1.15)

where $\text{Emb}[\phi]$ and $\text{Dir}[f]$ are the minimums over the respective homotopy classes (Theorem 6). Since Dirichlet energy can be interpreted as the elastic energy of a stretched rubber band network, $\text{Emb}[\phi] < 1$ can therefore also be interpreted as saying that $\Gamma_1$ is “looser” than $\Gamma_2$, however the two rubber band networks are stretched. See Remark 6.30.

Other examples are in Propositions 5.18 and 5.19. This gives a kind of duality between curves in an elastic graph $\Gamma$ and maps from $\Gamma$ to length graphs. If we think of curves as living in a “vector space” and maps to length graphs in its “dual”, then the embedding energy can be interpreted as an “operator norm”.

This theory of conformal graphs is largely parallel to the theory of conformal (Riemann) surfaces with boundary, where we again have a number of energies:

$$\ell^2$$

(1.16)

Again, each arrow is marked by the appropriate energy for measuring that type of map. $\text{Dir}$ and $\text{EL}$ are again Dirichlet energy and extremal length, but on surfaces rather than graphs. $\text{SF}$ is new; it is the stretch factor of a homotopy class of a topological embedding $[\phi]: \Sigma_1 \to \Sigma_2$ between Riemann surfaces. In general, we do not know a direct expression analogous to Equation (1.2), so $\text{SF}$ is defined to be the minimal ratio of extremal lengths (Definition 6.2, analogous to Equation (1.15)). When there is no conformal embedding of $\Sigma_1$ in $\Sigma_2$ in the given homotopy class, there
is a direct expression: $SF[\phi]$ is given by the minimal quasi-conformal constant in the homotopy class (Theorem 3).

**Remark 1.17.** To prove Theorem 1, we actually do not need to consider length graphs or the Dirichlet energy at all. (The proofs go through extremal length instead.) However, they illuminate the overall structure. In particular, it is not clear why one would consider elastic graphs without the rubber-band motivation.

**Remark 1.18.** The appearance of length squared in (1.6) and (1.16) is easy to justify on the grounds of units. Extremal length itself behaves like the square of a length, in the sense that if we take $k$ parallel copies of a curve, the extremal length multiplies by $k^2$. Likewise, if the lengths on the target of a harmonic map are multiplied by $k$, the harmonic map remains harmonic while the Dirichlet energy is multiplied by $k^2$.

1.3. **History and prior work.** Although Equation (1.2) appears to be new, Jeremy Kahn’s notion of domination of weighted arc diagrams [Kah06] is essentially equivalent. See Section 8.5.

Theorem 1 is closely related to Theorem 8.4, which characterizes when a rational map exists in terms of conformal embeddings of surfaces. Theorem 8.4 has been a folk theorem in the community for some time.

For polynomials, Theorem 1 reduces to a previously-known characterization in terms of expansion on the Hubbard tree; see Theorem 8.12 in Section 8.3.

1.4. **Organization.** After a section giving some examples of how to apply Theorem 1, this paper is organized by topics moving up a dynamical hierarchy. For each topic we give first the conformal surface notions and then the graph notions.

- We start with notions depending only on a single conformal surface or graph: *Dirichlet* (or *rubber band*) *energy* of (harmonic) maps (Section 4) and *extremal length* of curves (Section 5).
- Next come notions depending on a map between surfaces or graphs. This is the *stretch factor* or *embedding energy*, which generalizes the Teichmüller distance and characterizes conformal embeddings of surfaces (Section 6).
- Next is the dynamical theory of iterated maps (Section 7). Here we find another number, the *asymptotic stretch factor*, which characterizes rational maps (Section 8).

The paper is organized by a logical hierarchy, rather than what is necessarily pedagogically best; the reader is encouraged to skip around.

1.5. **Acknowledgements.** I would like to thank Matt Bainbridge, Steven Gortler, Richard Kenyon, Sarah Koch, Tan Lei, and Dan Margalit for many helpful conversations. Maxime Fortier Bourque pointed me towards Ioffe’s theorem [Iof75] and had numerous other insights.

This project grew out of extensive conversations with Kevin Pilgrim, who helped shape my understanding of the subject in many ways. Many of the arguments were developed jointly with him. Notably, he communicated Theorem 8.4 to me, and Theorem 4 is joint work with him. Theorem 5 is joint work with Jeremy Kahn, who also contributed substantially throughout.

Above all, I would like to thank William Thurston for introducing me to the subject and insisting on understanding deeply.
2. Examples

Here, we give some more substantial examples of Theorem 1.

2.1. Polynomials: The rabbit and the basilica. Theorem 1 is not very interesting for polynomials, as every topological polynomial with a branch-point in each cycle is equivalent to a polynomial. The extension of Theorem 1 to the general topological polynomial case is somewhat more interesting, but is equivalent to known results on expansion on the Hubbard tree; see Section 8.3. Nevertheless, we will look at some examples, both to see what the stretch factors are and to use them for matings.

Example 2.1. We first look at the "rabbit" polynomial, the post-critically finite polynomial \( f_1(z) = z^2 + c \) with \( c \approx -0.1226 + 0.7449i \). The critical point moves in a 3-cycle

\[
\begin{array}{c}
\vdots \quad c^2 + c \quad c \quad 0 \\
(2) \quad (0) \quad (2)
\end{array}
\]

The optimal elastic graph \( \Gamma_1 \) and its cover \( \tilde{\Gamma}_1 \) are

\[
\begin{array}{c}
\tilde{\Gamma}_1 \\
\Gamma_1
\end{array}
\]

with \( \text{SF}[\tilde{\Gamma}_1 \to \Gamma_1] = 2^{-1/3} \), which is less than one, as expected. (There is another marked point at infinity, not shown.)

Example 2.2. Another graph \( \Gamma_2 \) that works to prove that the rabbit polynomial is realizable is

\[
\begin{array}{c}
\tilde{\Gamma}_2 \\
\Gamma_2
\end{array}
\]

The black edges have the indicated lengths, which come from looking at the external rays landing at the \( \alpha \) fixed point of \( f_1 \). Give the colored edge an equal and long elastic length (say, 100). There is a natural map \( \phi_0 : \tilde{\Gamma}_2 \to \Gamma_2 \) as follows.

- The outside circle is mapped to the outside circle, with derivative \( |\phi_0'| = 1/2 \).
- The colored segments on the lower right of \( \tilde{\Gamma}_2 \) is squashed out to the lower-right boundary of \( \Gamma_2 \), with derivative \( |\phi_0'| \) on order of 1/100. Thus \( \text{Fill}_{\phi_0} \approx 1/2 \) on the corresponding portion of \( \Gamma_2 \).
- The colored segments in the upper left of \( \tilde{\Gamma}_2 \) are mapped to the colored segments in \( \Gamma_2 \), with \( |\phi_0'| = 1 \).
Because of the last point, this map has $\text{Emb}[\phi_0] = 1$, which is not good enough. Let $\phi_1$ be the result of “pulling in” very slightly the image of the ends of the upper-left colored segments of $\Gamma_2$, so they map to the interior of the colored segments of $\Gamma_2$. This decreases the derivative on the colored segments to less than one, while only increasing the derivative on the outside circle slightly; thus $\text{Emb}[\phi_1]$ is very slightly less than 1, as desired.

**Example 2.4.** We can perform a similar trick for other polynomials. For instance, the basilica polynomial $f_3(z) = z^2 - 1$ has a spine $\Gamma_3$, with cover given by

\begin{equation}
\begin{array}{c}
\text{\includegraphics[width=0.7\textwidth]{example24.png}}
\end{array}
\end{equation}

The same argument as above (giving the purple edge a long length, and pushing the right purple edge out to the boundary) shows that there is a map $\phi: \Gamma_3 \rightarrow \Gamma_3$ with $\text{Emb}[\phi] < 1$. (In this case the optimal stretch factor is $2^{-1/2}$ and is not realized with a graph $\Gamma_3$ with this topology).

### 2.2. Matings

We can use the techniques of Section 2.1 to show that some matings of polynomials are geometrically realizable.

**Example 2.6.** We can glue together the figures in Equations (2.3) and (2.5):

\begin{equation}
\begin{array}{c}
\text{\includegraphics[width=0.7\textwidth]{example26.png}}
\end{array}
\end{equation}

This gives a graph spine $\Gamma_4$ and cover $\Gamma_4$ representing the formal mating of the rabbit and the basilica. We can find a map $\phi: \Gamma_4 \rightarrow \Gamma_4$ with embedding energy less than 1:

- Assign the black mating circle in $\Gamma_4$ total length 1, divided according to the angles of the external rays.
- Give all colored edges an equal and large length.
- Pull this metric back to $\Gamma_4$, and map $\Gamma_4$ to $\Gamma_4$ by pushing colored arcs out to the black circle as in Examples 2.2 and 2.4.
- Pull the map in slightly where colored vertices meet the black circle.

The result has $\text{Fill}_\phi(y)$ slightly larger than 1/2 when $y$ is on the black circle, and $\text{Fill}_\phi(y)$ slightly less than one on the colored edges.

Naturally, the technique of Example 2.6 cannot always work, as sometimes the mating is not geometrically realizable.
Example 2.8. If we try to mate a basilica with a basilica, we get these graphs:

\[ (2.9) \]

\[ \begin{array}{c}
\hat{E} & \hat{B} & \hat{A} \\
\downarrow \hat{D} & \downarrow \hat{D}
\end{array} \]

If we try to use the same technique as before, it doesn’t work, as there are two points on the black circle that we attempt to pull in two different directions. Indeed, the left green-purple circle is mapped to a green-purple circle, so must have derivative at least 1: it is an obstruction to the mating. (In this case, it is a Lévy cycle).

2.3. Slit maps. Given a branched self-cover \( f : (S^2, \partial) \to (S^2, \partial) \) and an arc \( A \) with endpoints in \( P \), there is a blowing up construction which produces a map \( f_A \) that agrees with \( f \) outside of a neighborhood of \( A \) and maps that neighborhood surjectively on to \( S^2 \). Pilgrim and Tan Lei showed that, if \( f \) is a rational map, these blow ups frequently are as well [PT98]. We will restrict attention to cases where the initial map \( f \) is the identity, in which case the theorem becomes the following.

Theorem 2.10. Let \( P \subset S^2 \) be a finite graph, and let \( G \subset S^2 \) be a finite embedded graph with endpoints on \( P \). Then \( \text{id}_G \) is a rational map iff \( G \) is connected.

We can give a new proof in the harder direction, when \( G \) is connected. We start with a simple example. From the connected planar graph

\[ G = \begin{array}{c}
\circ \\
\circ
\end{array} \]

take as spine \( \Gamma_4 \) the spherical dual to \( G \). Then \( \tilde{\Gamma}_4 \) is obtained by taking the connect sum of \( \Gamma_4 \) with four extra copies \( \Gamma_4^i \) of \( \Gamma_4 \), one for each edge:

For any metric on \( \Gamma_4 \), there is a natural map \( \phi : \tilde{\Gamma}_4 \to \Gamma_4 \) that maps most of each copy \( \Gamma_4^i \) to a point in the center of the corresponding edge of \( \Gamma_0 \). This map \( \phi \) has derivative equal to 1/2 or 0 everywhere. We have \( \text{Emb}(\phi) = 1/2 \), and consideration of the red edge shows that this is optimal.

This example generalizes immediately to a general connected graph \( G \), except that \( \text{Emb}[\phi] \) will only be 1/2 when \( G \) has a univalent vertex; otherwise, \( \text{Emb}[\phi] \) will be
strictly smaller. We have thus proved Theorem 2.10, with some additional information about the stretch factor.

2.4. Behavior under iteration. We can iterate Example 1.1:

(See Section 7.1 for details on iteration.) In this example, the embedding energy behaves well, in the sense that the $k$'th iterate $\phi_k$ is optimal for embedding energy (using the edge lengths and concrete initial map from Example 1.3):

$$\text{Emb} [\psi_k] = \text{Emb}(\psi_k) = (\sqrt{2})^{-k}.$$ 

Such good behavior is not generally the case.

Example 2.11. The map

$$f(z) = \frac{1 + z^2}{1 - z^2}$$

is represented combinatorially by a graph $\Gamma_5$ and maps $\pi, \phi: \tilde{\Gamma}_5 \Rightarrow \Gamma_5$:

A case analysis shows that for any metric on $\Gamma$ and any map $\psi \in [\phi]$ with $\text{Emb}(\psi) < 1$, there is some power $1 \leq k \leq 4$ so that the $k$-fold iterate $\psi_k$ has local back-tracking. This implies that

$$\text{Emb}[\psi_4] < \text{Emb}[\psi]^4,$$

regardless of the initial metric. Similar facts hold for any of the other three graphs homotopy equivalent to $\Gamma_5$. Thus, regardless of the choice of initial spine, we have $\text{SF} [\phi] < \text{SF} [\phi]$. (See Section 7.2 for definitions.)

Table 1 gives a sample of experimental data for this map. By comparison, $\overline{\text{SF}} [\phi]$ is $2^{-1/3} \approx 0.793701$. 

| α  | β  | γ  | Emb[$f^{(n)}$]^{1/n} |
|----|----|----|---------------------|
| 1  | 1  | 1  | 1.500000 0.879465 0.949914 0.852017 0.873482 |
| 4  | 1  | 4  | 0.864357 0.839121 0.819495 0.810462 0.807620 |
| 4.5| 1  | 3.7| (3/4)^{1/2} (5/8)^{1/3} (23/48)^{1/4} (187/480)^{1/5} |
| 1  | 0  | 1  | = 1.000 ≈ 0.866 ≈ 0.855 ≈ 0.832 ≈ 0.828 |

Table 1. Results of computation for Example 2.11. The numbers $\alpha$, $\beta$, and $\gamma$ are the lengths of the blue, red, and green edges, respectively. The lengths (4.5, 1, 3.7) are close to optimal for the first iterate.
3. Setting

3.1. Surfaces. We work with compact, oriented surfaces $\Sigma$ with boundary. It is sometimes convenient to think about the double $D\Sigma = \Sigma \cup_B \overline{\Sigma}$, which has no boundary.

Definition 3.1. A curve on $\Sigma$ is an immersion of a 1-manifold with boundary in $\Sigma$, with boundary mapped to the boundary. The 1-manifold need not be connected; if it is, the curve is said to be connected. Curves are considered up to homotopy within the space of all maps taking the boundary to the boundary (not necessarily immersions). A curve is simple if it is embedded (has no crossings). An arc is an interval component of a curve, and a loop is a circle component. A curve of type $+$ is a curve with only loop components and a curve of type $-$ is a curve with no loops parallel to the boundary. The geometric intersection number $i([C_1], [C_2])$ of two curves is the minimal number of intersections (without signs) between representatives of the homotopy classes $[C_1]$ and $[C_2]$.

A weighted curve is a positive linear combination of curves, where two parallel components may be merged and their weights added. The space of weighted simple curves on $\Sigma$ is denoted $c(\Sigma)$. If $\Sigma$ has boundary, then we distinguish two subsets:

- $C^+(\Sigma)$ is the space of weighted simple curves of type $+$; and
- $C^-\Sigma)$ is the space of weighted simple curves of type $-$.

Remark 3.2. As curves need not be connected, they are what other authors would call a multi-curve.

There are several different geometric structures one can put on a surface. First, we can consider a conformal or complex structure $\omega$ on $\Sigma$, considered up to isotopy.

The next two structures deal with measured foliations (or equivalently measured laminations). We always consider measured foliations up to homotopy and Whitehead equivalence. Given a measured foliation $F$ and a curve $C$, we can compute $i([C], F)$, the minimal (transverse) length of any curve isotopic to $C$ with respect to $F$. This is unchanged under Whitehead equivalence, and the converse is true: two measured foliations are Whitehead equivalent iff the transverse lengths of all curves within an appropriate dual class (as specified below) are the same.

For a surface $\Sigma$ with no boundary, there is only one type of measured foliation, and they form a finite-dimensional space $MF(\Sigma)$. The space of curves $c(\Sigma)$ is dense in $MF(\Sigma)$, and measuring lengths on $c(\Sigma)$ gives an embedding $i: MF(\Sigma) \hookrightarrow \mathbb{R}^{c(\Sigma)}$.

On a surface with boundary, measured foliations come in two natural flavors. $MF^+(\Sigma)$ is the space of foliations that are parallel to the boundary (i.e., so the transverse length of the boundary is 0). $C^+(\Sigma)$ is dense in $MF^+(\Sigma)$, and measuring lengths of curves in $C^+(\Sigma)$ gives an embedding $i: MF^+(\Sigma) \hookrightarrow \mathbb{R}^{C^-(\Sigma)}$.

Dually, $MF^-(\Sigma)$ is the space of measured foliations without boundary annuli. It is the closure of $C^-(\Sigma)$, and there is an embedding $i: MF^-(\Sigma) \hookrightarrow C^+(\Sigma)$.

For a closed surface $\Sigma$, we define $MF^+(\Sigma)$ and $MF^-(\Sigma)$ to be equal to $MF(\Sigma)$.

Warning 3.3. The set of connected curves is not dense in $MF^{\pm}(\Sigma)$ if $\Sigma$ has non-empty boundary.

Finally, we consider quadratic differentials, in the following variants. A quadratic differential on $\Sigma$ is locally of the form $q(z) (dz)^2$ with $q$ holomorphic, and determines
a half-turn surface structure on $\Sigma$, away from a finite number of singular cone points. (A half-turn surface is a surface with a chart where the overlap maps are translations or rotations by $\pi$.) $Q(\Sigma, \omega)$ is the space of quadratic differentials with finite area.

If $\Sigma$ has boundary, then $Q^2(\Sigma, \omega)$ is the space of quadratic differentials that are real on the boundary; this is isomorphic to the space of quadratic differentials on $(D\Sigma, D\omega)$ that are invariant with respect to the involution. We are most interested in a further subspace. For a Riemann surface $(\Sigma, \omega)$ with boundary, $Q^+(\Sigma, \omega)$ is the subset of $Q^2(\Sigma, \omega)$ that is non-negative everywhere on each boundary component, or equivalently where the boundary is horizontal (rather than vertical). For a surface $\Sigma$ with no conformal structure, $Q^+(\Sigma)$ is the space of pairs of a conformal structure $\omega$ on $\Sigma$ and a quadratic differential in $Q^+(\Sigma, \omega)$, considered up to isotopy. Equivalently, a point in $Q^+(\Sigma)$ is a half-turn surface structure on $\Sigma$ with horizontal boundary.

From a quadratic differential $q \in Q^+(\Sigma)$, we can get horizontal and vertical measured foliations

$$q^h \in MF^+(\Sigma), \quad q^v \in MF^-(\Sigma).$$

At a point $x \in \Sigma$, the vectors $v \in T_x \Sigma$ that are tangent to one of these measured foliations are those for which $q(v, v) \geq 0$ for $q^h$ or $q(v, v) \leq 0$ for $q^v$. The transverse measure of an arc $\gamma(t)$ is given by

$$q^h(\gamma) := \int |\operatorname{Im} \sqrt{q(\gamma'(t), \gamma'(t))}| \, dt \quad q^v(\gamma) := \int |\operatorname{Re} \sqrt{q(\gamma'(t), \gamma'(t))}| \, dt.$$

A quadratic differential is more or less the combination of two of the other three types of data.

- A conformal structure and a measured foliation $F^+ \in MF^+(\Sigma)$ uniquely determines a quadratic differential in $q \in Q^+$ with $F^+ = q^h$. This is the heights theorem [HM79, Ker80, Gar84, MS84, Gar87].
- Let $F^+, F^-$ be a pair of measured foliations in $MF^+(\Sigma)$ and $MF^-(\Sigma)$, respectively. Then generically there is a unique half-turn surface structure in $q \in Q^+(\Sigma)$ with the given foliations as horizontal and vertical foliations, respectively [GM91, Theorem 3.1].
- On the other hand, given a conformal structure and measured foliation $F^- \in MF^-(\Sigma)$, there is not always a quadratic differential in $Q^+$ with $F^-$ as its vertical measured foliation. You can always double the situation and consider the foliation $D(F^-)$ on $D(\Sigma)$, which has an associated quadratic differential restricting to a unique $q \in Q^2(\Sigma)$, but $q$ need not be in $Q^+(\Sigma)$; portions of $\partial \Sigma$ might be vertical rather than horizontal. However, if there is such a quadratic differential, it is unique, as the doubling argument shows.

The types of structure on surfaces are summarized on the top row of Table 2.

### 3.2. Convexity

One key fact, used in Section 5.1, is that there is a natural convex structure on $MF^\pm(\Sigma)$. Recall that there are (several) natural coordinates for measured foliations.

- For $MF^+(\Sigma)$ of a surface with boundary, pick a maximal collection of non-parallel disjoint simple arcs on $\Sigma$, and measure the transverse lengths of each.
For \( \mathcal{MF}(\Sigma) \) of a closed surface or \( \mathcal{MF}^-(\Sigma) \) of a surface with boundary, take Dehn-Thurston coordinates with respect to some marked pair-of-pants decomposition of \( \Sigma \). Normalize the Dehn-Thurston twist parameter so that twist 0 corresponds to measured foliations that are invariant under reversing the orientation of \( \Sigma \). (See, e.g., [Thu].)

Call any of these coordinate systems *canonical coordinates*.

**Definition 3.4.** A function on \( \mathcal{MF}^\pm(\Sigma) \) is *strongly convex* if it is convex as a function on \( \mathbb{R}^k \) for each choice of canonical coordinates.

This definition appears quite restrictive, since there are infinitely many different canonical coordinates. However, such functions do exist.

**Theorem 2.** For \( b \in \{+, -\} \), let \( f \) be a function on weighted curves of type \( b \) so that

- \( f \) does not increase under smoothing of essential crossings:

\[
f\left(\begin{array}{c}
\ominus \\
\ominus
\end{array}\right) \geq f\left(\begin{array}{c}
\ominus \\
\ominus
\end{array}\right),
\]

where the crossing is essential and the strands have the same weight, and

- \( f \) is convex under union: for \( C_1 \) and \( C_2 \) two curves,

\[
f\left(\left(C_1 \cup C_2\right)/2\right) \leq \left(f[C_1] + f[C_2]\right)/2.
\]

Then \( f \) extends uniquely to a continuous, strongly convex function on \( \mathcal{MF}^b(\Sigma) \).

An essential crossing of a curve is (somewhat loosely) one that cannot be removed by homotopy. Note that \( C_1 \cup C_2 \) need not be simple, even if \( C_1 \) and \( C_2 \) are.

**Example 3.7.** Take any geodesic metric \( g \) on \( \Sigma \), and let \( \ell_g[C] \) be the minimal length of a curve in the homotopy class \([C]\) with respect to the metric \( g \). Then Equation (3.5) is true, as smoothing crossings in the geodesic representative can only decrease length, and Equation (3.6) is true by definition (with equality). Thus all of these functions extend to continuous functions on \( \mathcal{MF}^b(\Sigma) \). Special cases of interest include when

- the metric is a hyperbolic metric on \( \Sigma \), or
- the metric degenerates so that the lengths approach the transverse measure with respect to a measured foliation.

As an example of this degeneration, the function \( F \mapsto \iota([C], F) \) for a fixed background curve \( C \) is a strongly convex function on \( \mathcal{MF}^b(\Sigma) \).
Lemmas 5.5 and 5.6 below say that extremal length also satisfies the same conditions.

Proof sketch of Theorem 2. Fix a canonical set of coordinates on $\mathcal{MF}^b(\Sigma)$. Given two rational measured foliations $F_0, F_1 \in \mathcal{MF}^b$, let $F_{1/2} = (F_0 \oplus F_1)/2$ be the midpoint of the straight-line path between them, where $\oplus$ means adding the chosen canonical coordinates. (This depends on the coordinate system.) Since they are rational, $F_0$ and $F_1$ are represented by weighted simple curves, as is $F_{1/2}$. Analysis of the coordinates shows that $F_{1/2}$ is obtained from $(F_0 \cup F_1)/2$ by smoothing crossings (for some choice of resolution of the crossings). Thus, by Equations (3.5) and (3.6),

$$f[F_{1/2}] \leq f[(F_0 \cup F_1)/2] \leq (f[F_0] + f[F_1])/2,$$

which implies that $f$, when defined, is a convex function. But $f$ is defined on the dyadic rational points in $\mathcal{MF}^b(\Sigma)$, and continuity of $f$ follows from convexity [Roc70, Theorem 10.1].

3.3. Graphs. In this paper, a graph $\Gamma$ is a connected 1-complex, with possibly multiple edges and self-loops. A ribbon graph has in addition a cyclic ordering on the ends incident to each vertex; this gives a canonical thickening of $\Gamma$ into an oriented surface with boundary $N\Gamma$, as in Figure 2.

A curve on a graph $\Gamma$ is a map from a 1-manifold $C$ into $\Gamma$, considered up to homotopy. We do not admit arcs here, and up to homotopy we can assume that curves are taut: they do not backtrack on themselves.

We can put geometric structures on graphs corresponding loosely to the four geometric structures on surfaces above, as outlined in Table 2.

First, corresponding to a conformal structure, we can consider an elastic graph $\Gamma$, with an (elastic) weight on the edges: for each edge $e$ of $\Gamma$, give a positive measure which is absolutely continuous with respect to the Lebesgue measure. Up to equivalence, we effectively just give a total measure $\alpha(e)$ on each edge.

There are at least two ways to interpret this measure $\alpha$. On one hand, we can create a rubber-band network. For each edge $e$, take an idealized rubber band with spring constant $1/\alpha(e)$, so the Hooke’s Law energy when the edge is stretched to length $\ell(e)$ is

$$\text{Energy}(e) = \ell(e)^2/2\alpha(e).$$

Attach these rubber bands at the vertices. Note that, as for real rubber bands, a longer section of rubber band (larger $\alpha(e)$) is easier to stretch. Unlike for real rubber bands, the resting length is 0.

Alternately, we can think of $\alpha(e)$ as defining a family of rectangle surfaces: given elastic weights on a ribbon graph $\Gamma$ and a constant $\epsilon > 0$, define a conformal surface
Figure 3. Geometrically thickening a graph. Left: An edge of length $\alpha$ is thickened to an $\alpha \times \epsilon$ rectangle. Right: At a vertex, glue each half of the end of each rectangle to one of the neighbors. This is a more geometrically precise version of Figure 2.

$N_c\Gamma$ by taking a rectangle of size $\alpha(e) \times \epsilon$ for each edge $e$ and gluing them according to the ribbon structure at the vertices, as indicated in Figure 3. $\Gamma$ should be thought of as the limit of the Riemann surfaces $N_c\Gamma$ as $\epsilon \to 0$.

The next geometric structure is a width graph, which is a graph $C$ with a width on the edges: for each edge $e$, there is a width $w(e)$ satisfying a triangle inequality at each vertex $v$, as follows. Let $e_1, \ldots, e_n$ be the edges incident to $v$ (with an edge appearing twice if it is a self-loop). Then we require, for each $1 \leq j \leq n$,

$$w(e_j) \leq \sum_{1 \leq i \leq n \atop i \neq j} w(e_i).$$

(3.9)

$W(\Gamma)$ is the space of possible widths on an abstract graph $\Gamma$.

For a ribbon graph $\Gamma$, there is a natural surjective map $W: \mathcal{MF}^+(N\Gamma) \to W(\Gamma)$. For $e$ an edge of $\Gamma$, let $e^*$ be the arc on $N\Gamma$ dual to $e$. Then $W(F^+)$ takes the value $F^+(e^*)$ on the edge $e$. When $\Gamma$ is trivalent (i.e., the vertices of $\Gamma$ have degree $\leq 3$), $W$ is an isomorphism. In general there is ambiguity about how to glue at the vertices.

Dually, a length graph $K$ is a graph together with an assignment $\ell$ of a non-negative length to each edge of $\Gamma$, which we think of as defining a pseudo-metric on $\Gamma$. (It is a pseudo-metric because points can be distance zero.) $\mathcal{L}(\Gamma)$ is the space of lengths on $\Gamma$.

If $\Gamma$ is a ribbon graph, there is a natural map $L^*: \mathcal{L}(\Gamma) \to \mathcal{MF}^-(N\Gamma)$, defined by

$$L^*(\ell) = \sum_{e \in \text{Edge}(\Gamma)} \ell(e) \cdot e^*$$

for $\ell$ a loop on $\Gamma$. This map is not surjective, but if $\Gamma$ is trivalent, $L^*$ maps onto an open subset of $\mathcal{MF}^-(N\Gamma)$.

Finally, a strip graph $S$ is a graph in which each edge $e$ is equipped with
- a length $\ell(e)$,
- a width $w(e)$, and
- an elastic weight (aspect ratio) $\alpha(e)$,

so that $\ell(e) = \alpha(e)w(e)$ and the widths satisfying the triangle inequalities. Equation (3.9). A strip graph has a natural underlying elastic graph $\Gamma(S)$, width graph
C(S), and length graph K(S). It also has a total area

\[(3.10) \quad \text{Area}(S) = \sum_{e \in \text{Edge}(S)} \ell(e)w(e) = \sum_{e \in \text{Edge}(S)} \frac{\ell(e)^2}{\alpha(e)} = \sum_{e \in \text{Edge}(S)} \alpha(e)w(e)^2.\]

(Compare this to elastic energy, Equation (3.8) above, and extremal length on graphs, Equation (5.13) below.)

**Warning 3.11.** Be careful to distinguish between lengths and elastic weights. They are determined by the same data (a measure or equivalently a length on each edge), but they are interpreted differently. If the elastic weights are interpreted as defining a rubber band network, then the lengths can be interpreted as lengths of a system of pipes through which the rubber bands can be stretched. Alternatively, when the elastic weights are aspect ratios of rectangles (length/width), the lengths are just the length of the rectangles.

As for surfaces, two of the three other types of data determine a strip graph, except that a choice of elastic weights and lengths may not correspond to a strip graph, as the triangle inequalities on the widths may be violated.

**Remark 3.12.** In the thickening of an elastic graph \(\Gamma\) into a surface \(N, \Gamma\) in Figure 3, the precise details of how you glue the rectangles at the vertices are irrelevant in the limit, in the following sense. If we pick two different ways of gluing at a vertex (e.g., gluing different proportions to the left and right) and get two families of surface \(N^1, \Gamma\) and \(N^2, \Gamma\), then in the limit as \(\epsilon \to 0\) the minimal quasi-conformal constant of maps between \(N^1, \Gamma\) and \(N^2, \Gamma\) goes to 1.

In particular, a strip graph \(S\) has a thickening \(N, S\), where each edge \(e\) is replaced by a rectangle of size \(\ell(e) \times \epsilon w(e)\), and the rectangles are glued in a width-preserving way at the vertices. (This gluing is possible by the triangle inequalities.) If \(\Gamma(S)\) is the underlying elastic graph of \(S\), then for \(\epsilon \ll 1\) the surfaces \(N, S\) and \(N, \Gamma(S)\) are nearly conformally equivalent. See Proposition 5.15 for some related estimates.

**3.4. Extensions.** The graphs \(\Gamma\) we are considering in this paper are usually topological spines of the corresponding surface \(N, \Gamma\) (i.e., \(N, \Gamma\) deformation retracts onto \(\Gamma\)). More generally, we may consider a graph \(\Gamma\) and surface \(\Sigma\) with a \(\pi_1\)-surjective embedding \(G \hookrightarrow \Sigma\), or more generally still a graph \(\Gamma\), group \(G\), and surjective map \(\pi_1(\Gamma) \to G\) (i.e., a generating graph for \(G\)). Much of the theory extends to this case, as we will comment as we go along.

It is also sometimes convenient to generalize from surfaces to orbifolds. In the setting of groups, this means considering maps from \(\pi_1(\Gamma)\) to the orbifold fundamental group \(\pi_1^{\text{orb}}(\Sigma)\). In the setting of graphs, we consider graphs with marked points that are required to be mapped to marked (orbifold) points. In particular, this is required for a proper statement of Theorem 1 if we drop the restriction that there be a branch point in each cycle of \(P\).

**4. Harmonic maps and Dirichlet energy**

We now turn to harmonic maps, maps that minimize some form of Dirichlet energy.
4.1. **Harmonic maps from surfaces.** Given a conformal Riemann surface \((\Sigma, \omega)\), a length graph \((K, \ell)\), and a Lipshitz map \(f: \Sigma \to K\), the *Dirichlet energy* of \(f\) is

\[
\text{Dir}(f) := \int_{\Sigma} |\nabla f(z)|^2 \, dA.
\]

Here, we have picked an (arbitrary) Riemannian metric \(g\) in the given conformal class, \(dA\) is the area measure with respect to \(g\), and \(|\nabla f(z)|\) defined to be the best Lipshitz constant of \(f\) at \(z\) with respect to \(g\) and \(\ell\). (This agrees with the norm of the ordinary gradient when \(f(z)\) is in the interior of an edge of \(K\) and \(f\) is differentiable.)

We also define \(\text{Dir}[f]\) to be the lowest energy in the homotopy class of \(f\):

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

This optimum is achieved [EF01, Theorem 11.1]. The optimizing functions are called *harmonic maps* and define canonical quadratic differentials in \(Q^+(\Sigma)\).

**Warning** 4.2. The connection between harmonic maps and quadratic differentials is not as tight as you might expect. Given a closed surface \(\Sigma\) and a weighted simple curve \(C\), we can construct a length graph \(K(C)\), with vertices the connected components of \(\Sigma \cap C\) and edges given by components of \(C\), with lengths of edges of \(K(C)\) given by weights on \(C\). There is a natural homotopy class of maps \(\Sigma \to K(C)\). The Dirichlet minimizer in this homotopy class sometimes, but not always, recovers the Jenkins-Strebel quadratic differential with vertical foliation given by \(C\).

A similar construction using maps from the universal cover of \(\Sigma\) to an \(\mathbb{R}\)-tree does recover the quadratic differential with given vertical foliation. Wolf used this to give an alternate proof of the Heights Theorem [Wol98].

Alternatively, we could generalize the target space, and allow \(K\) to be a general non-positively curved polyhedral complex. In this context, we can take a complex \(K_s(C)\) made of very thin tubes around the edges around the graph \(K(C)\) defined above. Then there is a closer connection between harmonic maps to \(K_s(C)\) and the Jenkins-Strebel quadratic differential. We do not pursue this further here.

4.2. **Harmonic maps from graphs.** Given an elastic graph \((\Gamma, \alpha)\) and a length graph \((K, \ell)\), the *Dirichlet energy* of a Lipshitz map \(f: \Gamma \to K\) is

\[
\text{Dir}(f) := \int_{\Gamma} |f'(x)|^2 \, dx = \int_K \text{Fill}_f(y) \, dy
\]

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

Here, \(|f'(x)|\) is the derivative of \(f\) with respect to the natural coordinates given by \(\alpha\) on \(\Gamma\) and \(\ell\) on \(K\). This derivative is not defined at vertices, but these points are negligible. Alternatively, \(|f'(x)|\) is the best Lipshitz constant of \(f\) at \(x\). \(\text{Fill}_f(y)\) is the *filling function* of \(f\) at \(y \in K\), the sum of derivatives at preimages:

\[
\text{Fill}_f(y) := \sum_{f(x) = y} |f'(x)|.
\]

The two expressions for \(\text{Dir}(f)\) are related by an easy change of variables.

Minimizers of Dirichlet energy are harmonic functions in the following sense.
Definition 4.6. A function $f : \Gamma \to K$ from an elastic graph $\Gamma$ to a length graph $K$ is harmonic if the following conditions are satisfied.

1. The map $f$ is piecewise linear.
2. The map $f$ does not backtrack (i.e., $f$ is locally injective on each edge of $\Gamma$).
3. The derivative $|f'(x)|$ is constant on the edges of $\Gamma$ when defined. As a result, for $e$ an edge of $\Gamma$ we may write $|f'(e)|$ for the common value at any point on the edge.
4. If a vertex $v$ of $\Gamma$ maps to the interior of an edge of $K$, with edges $e_1, \ldots, e_k$ incident on the left and edges $e_{k+1}, \ldots, e_{k+l}$ incident on the right, then

$$\sum_{i=1}^{k} |f'(e_i)| = \sum_{i=1}^{l} |f'(e_{i+k})|.$$  

In particular, for each edge $E$ of $K$, the filling function $\text{Fill}_f(y)$ is constant on $E$, and we may write $\text{Fill}_f(E)$.

5. If a vertex $v$ of $\Gamma$ maps to a vertex $w$ of $K$, let $E_1, \ldots, E_k$ be the germs of edges of $K$ incident to $w$. Then we have a vertex balancing condition: for $i = 1, \ldots, k$,

$$\sum_{e \text{ incident to } v} |f'(e)| \leq \sum_{1 \leq j \leq k} \sum_{e \in E_j} |f'(e)|.$$  

Here, the notation “$\sum_{e \in E_i}$” means the sum over all germs of edges of $\Gamma$ that locally map to $E_i$.

Theorem 4.9. A function $f : \Gamma \to K$ is a local minimum for the Dirichlet energy within its homotopy class iff it is harmonic. Every local minimum is also a global minimum.

From the rubber bands point of view, the intuition behind Definition 4.6 is that $|f'(x)|$ is the tension in the rubber band, which is constant along the edge. The net force on a vertex mapping to the interior of an edge must be zero; this is condition (4). For a vertex $v$ mapping to a vertex $w$, the net force pulling $v$ in any one direction away from $w$ cannot be too large; this is condition (5). Condition (4) can be thought of as the special case of condition (5) when $w$ has only two incident edges.

Conditions (4) and (5) also imply that the derivatives $|f'(e)|$ form a valid width structure on $\Gamma$ (and that the filling functions $\text{Fill}_f(E)$ form a width structure on $K$). From the rectangular surface point of view, a harmonic map gives a tiling of $K$ with rectangles, with each edge $e$ of $\Gamma$ contributing a rectangle of aspect ration $\alpha(e)$. To illustrate this, consider the case of marked planar graphs. Let $\Gamma$ be a planar graph with two distinguished vertices $s$ and $t$ bordering the infinite face, and consider the minimizer of the Dirichlet energy among maps

$$(\Gamma, s, t) \rightarrow ([0, 1], \{0\}, \{1\})$$

mapping $\Gamma$ to the interval, taking $s$ to 0 and $t$ to 1. Brooks, Smith, Stone, and Tutte showed that this minimizer gives a rectangle packing of a rectangle [BSST40]. See Figure 4 for an example.

Remark 4.10. In the context of rectangle packings, it is traditional to use the language of resistor networks rather than elastic or rubber-band networks. The Dirichlet energy
in these two settings is the same for maps to an interval, but for general target graphs rubber bands are more flexible, because of the lack of orientations. Where rubber band optimizers are related to homotopy of maps from $\Gamma$ to a target graph and holomorphic quadratic differentials, electrical current or resistor network optimizers are related to homology of $\Gamma$ and to holomorphic differentials (not quadratic differentials).

4.3. **Relation to Lipshitz energy.** Another natural norm for maps $\phi: K_1 \to K_2$ between metric graphs is the *Lipshitz energy*

\[
\text{Lip}(\phi) := \text{ess sup}_{x \in K_1} |\phi'(x)|
\]

\[
\text{Lip}[\phi] := \inf_{\psi \in [\phi]} \text{Lip}(\psi).
\]

Compared to Equation (4.3), we take the $L^\infty$ norm of $|\phi'(x)|$ rather than the $L^2$ norm.

Let us note that

- $\text{Lip}(\phi) < 1$ iff $\phi$ is distance-decreasing, and
- the Lipshitz energy is *dynamical*, in the sense that it behaves well under composition:

\[
\text{Lip}[\phi \circ \psi] \leq \text{Lip}[\phi] \text{Lip}[\psi].
\]

The Lipshitz energy should be thought of as an invariant of a map between two length graphs, not as an invariant of a map from an elastic graph to a length graph.

It follows immediately from the definitions that Lip is sub-multiplicative with respect to both length and Dirichlet energy: for $C$ a curve on $K_1$ and $f: \Gamma \to K_1$ a map from an elastic graph,

\[
\ell_{K_2}[\phi \ast C] \leq \text{Lip}[\phi] \ell_{K_1}[C]
\]

\[
\text{Dir}[\phi \circ f] \leq \text{Lip}[\phi]^2 \text{Dir}[f].
\]
These are both special cases of Equation (1.13), and in both cases the inequalities are tight. For instance, we have [FM11, Proposition 3.11]

\begin{equation}
\text{Lip}[\phi] = \sup_{C \text{ curve on } K_1} \frac{\ell_{K_1}[\phi_* C]}{\ell_{K_1}[C]}.
\end{equation}

4.4. Computing Dirichlet energy. Given a homotopy class \([f]: \Gamma \to K\) of maps between an elastic graph \(\Gamma\) and a metric graph \(K\), how can one find a representative \(g \in [f]\) that minimizes Dirichlet energy? By Theorem 4.9, this is the same as finding a harmonic map in the homotopy class.

As with harmonic maps in other settings, this is easy to compute, with at least two reasonable approaches:

1. Repeated averaging, iteratively moving the image of each vertex of \(\Gamma\) to the weighted average of its neighbors in the universal cover \(\tilde{K}\) of \(K\). The average of a non-empty set \(S\) of points on a metric tree is the (unique) point \(x\) that minimizes the sum of squares of distances from \(x\) to points in \(S\). Note that this average can generically be on a vertex.

2. Linear or convex quadratic programming, based on the observation that once the combinatorics of the map are fixed, specifying which vertices of \(\Gamma\) go to which vertices or edges of \(K\), the Dirichlet energy is a convex, quadratic function of the positions along the edges, and thus the minimum can be found by solving linear equations. Here, one starts by guessing some combinatorics, and updating the combinatorics if it turns out not to be optimal.

The second approach should usually be faster, but also requires some more care in changing the combinatorics. The harmonic representative of \([f]\) is not unique in general; however, the set of harmonic representatives forms a convex set in a suitable sense.

5. Extremal length

5.1. Extremal length on surfaces. Given a Riemann surface \((\Sigma, \omega)\) and a curve \(C\) on \(\Sigma\), recall that the extremal length of \(C\) on \(\Sigma\) is (omitting analytic details)

\begin{equation}
\text{EL}_\omega[C] := \sup_{\rho: \Sigma \to \mathbb{R}_{>0}} \frac{\ell_{\rho g}[C]^2}{\text{Area}_{\rho g}(\Sigma)},
\end{equation}

where we make the following definitions.

- The metric \(g\) is an arbitrary metric in the conformal class \(\omega\).
- The metric \(\rho g\) is the metric \(g\) scaled by the conformal factor \(\rho\). (This may be a pseudo-metric.)
- The number \(\ell_{\rho g}[C]\) is the minimal length of any element of the homotopy class \([C]\) in the metric \(\rho g\).
- The number \(\text{Area}_{\rho g}(\Sigma)\) is the total area of \(\Sigma\) with respect to \(\rho g\).

Observe that scaling \(\rho\) by a global constant does not change the suprema.

We can extend Equation (5.1) to allow for weighted curves: for \(C = \sum_i a_i C_i\) a weighted curve, define \(\ell[C] := \sum_i a_i \ell[C_i]\), and use Equation (5.1) as before. (Note that this agrees with the earlier definition for integral weights.)

The following results are standard.
**Theorem 5.2** (Jenkins-Strebel). If \( C \) is a weighted simple curve, then the supremum in the definition of \( EL[C] \) is achieved, and the optimal metric is the metric on a half-turn surface associated to a quadratic differential with horizontal foliation equal to \( C \).

**Lemma 5.3.** Let \( \pi: \tilde{\Sigma} \to \Sigma \) be a covering map of degree \( d \). For \( C \) a weighted curve on \( \Sigma \), define \( \pi^{-1}C \) to be the inverse image of \( C \), with the same weights. Then \( EL[\pi^{-1}C] = d EL[C] \).

**Lemma 5.4.** If \( a \in \mathbb{R}_{>0} \) is a global weighting factor, then \( EL[aC] = a^2 EL[C] \).

We now turn to properties related to convexity, as in Section 3.2. The following two properties follow from elementary arguments.

**Lemma 5.5.** \( EL \) does not increase under smoothing of essential crossings:

\[
EL \left[ \begin{array}{c}
\mathcal{X}
\end{array} \right] \geq EL \left[ \begin{array}{c}
\mathcal{O}
\end{array} \right].
\]

Here, the two sides show a local picture of unweighted curves, or weighted curves with equal weights on the two local strands.

**Lemma 5.6.** If \( C_1 \) and \( C_2 \) are two weighted multi-curves, then

\[
EL[C_1 \cup C_2] \leq 2(EL[C_1] + EL[C_2]),
\]

where we keep the weights on each component of \( C_1 \) and \( C_2 \). More generally, for \( 0 \leq t \leq 1 \),

\[
EL[tC_1 \cup (1-t)C_2] \leq t EL[C_1] + (1-t) EL[C_2].
\]

**Remark 5.7.** In Lemma 5.6, we are taking the union of curves, not the union of path families, as sometimes appears in the theory of extremal length.

**Corollary 5.8.** \( EL \) extends uniquely to a continuous, strongly convex function on \( \mathcal{MF}^\pm(\Sigma) \).

**Proof.** Follows from Theorem 2. \( \square \)

**Remark 5.9.** Presumably the techniques in the proof of Corollary 5.8 can be extended to prove the rest of the Heights Theorem, in particular that the associated quadratic differential varies continuously as a function of the measured foliation.

5.2. **Extremal length on graphs.** By analogy with Equation (5.1), for \( C \) a weighted curve on an elastic graph \((\Gamma, \alpha)\), define

\[
(5.10) \quad EL_\alpha[C] := \sup_{\rho: \text{Edge}(\Gamma) \to \mathbb{R}_{\geq 0}} \frac{\ell_{\rho\alpha}[C]^2}{\text{Area}_{\rho\alpha}(\Gamma)},
\]

where

- The length metric \( \rho\alpha \) on \( \Gamma \) gives edge \( e \) the length \( \rho(e)\alpha(e) \).
- The number \( \ell_{\rho\alpha}[C] \) is the length of \( C \) with respect to \( \rho\alpha \), i.e.,

\[
(5.11) \quad \ell_{\rho\alpha}[C] := \sum_{e \in \text{Edge}(\Gamma)} n_C(e)\rho(e)\alpha(e),
\]

where \( n_C(e) \) is the weighted number of times that \( C \) runs over \( e \).
• Area\textsubscript{\(\rho a\)}(\(\Gamma\)) is the “area” of \(\Gamma\) with respect to \(\rho a\), defined to be
\begin{equation}
\text{Area}_{\rho a}(\Gamma) := \sum_{e \in \text{Edge}(\Gamma)} \rho(e)^2 \alpha(e).
\end{equation}

(The intuition is that each edge is turned into a rectangle of width proportional to \(\rho(e)\) and aspect ratio \(\alpha(e)\), and thus length \(\rho(e)\alpha(e)\).)

In fact the supremum in Equation (5.10) is easy to do. The optimum has \(\rho(e)\) proportional to \(n_{C}(e)\) and so
\begin{equation}
\text{EL}_{\alpha}[C] = \sum_{e \in \text{Edge}(\Gamma)} n_{C}(e)^2 \alpha(e).
\end{equation}

This formula extends immediately to a function on \(W(\Gamma)\), and satisfies Lemmas 5.3, 5.4, and 5.6. For Lemma 5.5, we need to pick a ribbon structure on \(\Gamma\) in order to make sense of “essential” crossings. With any such choice, Lemma 5.5 is true.

**Corollary 5.14.** Let \((\Gamma, \alpha)\) be an elastic spine for a surface \(\Sigma\). Then the function \(C \mapsto \text{EL}_{\alpha}[C]\) on \(C^{+}(\Sigma)\) extends uniquely to a strongly convex function on \(\mathcal{MF}^{+}(\Sigma)\).

5.3. **Relating graphs and surfaces.** We can now give a concrete relation between an elastic graph \((\Gamma, \alpha)\) and the associated family of conformal surfaces \(N_{\epsilon}\Gamma\). Write \(\text{EL}[C; \Gamma]\) for extremal length with respect to the elastic graph \(\Gamma\), and \(\text{EL}[C; \Sigma]\) for extremal length with respect to the conformal surface \(\Sigma\).

**Proposition 5.15.** Let \((\Gamma, \alpha)\) be an elastic ribbon graph with trivalent vertices, and let \(m = \min_{e} \alpha(e)\) be the smallest weight of any edge in \(\Gamma\). Then, for \(t < m/2\) and \(C\) any measured foliation on \(\Gamma\), we have
\[
\text{EL}[C; \Gamma] \leq t \text{EL}[C; N_{t}\Gamma] \leq \text{EL}[C; \Gamma] \cdot (1 + 8t/m).
\]

The proof involves finding, on the one hand, embeddings of sufficiently thick annuli into \(N_{t}\Gamma\), and, on the other hand, suitable test functions \(\rho\) on \(N_{t}\Gamma\) in Equation (5.1).

**Remark 5.16.** The restriction to trivalent graphs in Proposition 5.15 can presumably be removed. Note that the estimate depends only on the local geometry of \(\Gamma\), and thus is unchanged under covers.

5.4. **Duality with Dirichlet energy.** As mentioned earlier, extremal length is in some sense dual to Dirichlet energy. More precisely, we have the following.

**Proposition 5.17** (Sub-multiplicative). Let \(C\) be a curve in an elastic graph \(\Gamma\), and let \(f : \Gamma \rightarrow K\) be a harmonic map to a length graph. Then
\[
\ell[f_{*}C]^{2} \leq \text{Dir}[f] \text{EL}[C].
\]

**Proof sketch.** This is basically the definition of \(\text{EL}\) in Equation (5.10), noting the similarities between the area in Equation (5.12) and Dirichlet energy, Equation (4.3).

**Proposition 5.18** (Duality 1). Let \(f : \Gamma \rightarrow K\) be a harmonic map from an elastic graph to a length graph. Then there is a sequence of weighted curves \(C_{i}\) in \(\Gamma\) so that for all \(\epsilon > 0\) there is an \(i\) so that
\[
\ell[f_{*}C_{i}]^{2} \geq (1 - \epsilon) \text{Dir}[f] \text{EL}[C].
\]
Proof sketch. Take weighted curves $C_i$ so that $n_{C_i}(e)$ approximates $|f'(e)|$. □

**Proposition 5.19 (Duality 2).** Let $C$ be a curve in an elastic graph $\Gamma$. Then there is a length graph $K$ and a harmonic map $f : \Gamma \to K$ so that

$$\ell[f_* C]^2 = \text{Dir}[f] \text{EL}[C].$$

Proof sketch. Take $K$ to be $\Gamma$ with edge lengths $\alpha(e)n_{C_i}(e)$. □

The situation is less satisfactory for surfaces. Sub-multiplicativity in the sense of Proposition 5.17 is true (at least when the curve $C$ is embedded), as is Proposition 5.18. Issues related to Warning 4.2 make an analogue of Proposition 5.19 more delicate, although it is true if we allow non-positively curved polyhedral complexes as the target space.

**Remark 5.20.** The definition of extremal length on graphs in Equation (5.10) is somewhat backwards, in that it is a supremum of a ratio of energies over all metrics (or equivalently over all maps to length graphs). By analogy with Dirichlet energy (Equation (4.3)), it would be better to define the energy of a homotopy class as an infimum of some energy functional. Indeed, we could take Equation (5.13) as the primary definition.

This remark applies to extremal length on surfaces (Equation (5.1)) as well: it might be better to take a different definition as primary. Namely, recall that a conformal annulus $A$ has an extremal length $\text{EL}(A)$, which we may define as the inverse of the modulus. Then extremal length of a weighted multi-curve $C = \sum_i a_i C_i$ can be alternately defined as

$$\text{EL}[C] = \inf_{(A_i)} \sum_i a_i \text{EL}(A_i),$$

where the infimum runs over all disjoint embeddings of conformal annuli $A_i$ with core curves homotopic to $C_i$.

5.5. **Extensions.** There are several ways in which we can extend these notions of extremal length. First, we can consider graphs that are embedded in a surface, not necessarily as a spine.

**Definition 5.22.** A graph $\Gamma$ embedded in a surface $\Sigma$ is filling if each component of $\Sigma\setminus\Gamma$ is a disk or an annulus on the boundary of $\Sigma$.

**Definition 5.23.** Let $(K, \ell)$ a length graph with a filling embedding $\phi$ in a surface $\Sigma$. For $C$ a curve in $\Sigma$, define

$$\ell_K[C] := \inf_{\text{curve } \gamma \text{ on } \Gamma} \ell[\gamma].$$

Now for $(\Gamma, \alpha)$ a filling elastic graph in $\Sigma$, define $\text{EL}_{\Gamma, \alpha}[C]$ by

$$\text{EL}_{\Gamma, \alpha}[C] := \sup_{\rho : \text{Edge}(\Gamma) \to \mathbb{R}_{>0}} \frac{\ell_{\Gamma, \rho \alpha}[C]^2}{\text{Area}_{\rho \alpha}(\Gamma)}.$$

This is just like Equation (5.10), except that we consider homotopy classes in $\Sigma$ rather than in $\Gamma$. A similar notion of extremal length was considered by Duffin [Duf62], in the context of electrical networks and graphs with two marked points (as in Section 4.2).
The optimization in Definition 5.23 is no longer as easy, and the analogue of Equation (5.13) is more awkward to state. The result of the optimization gives a rectangular tiling of the surface with aspect ratios given by \( \alpha \), analogous to Figure 4.

More generally, we can consider graphs generating a group.

**Definition 5.24.** For \((\Gamma, \alpha)\) an elastic graph, \(\phi: \pi_1(\Gamma) \rightarrow G\) a surjective homomorphism onto a group \(G\), and \([g]\) a conjugacy class in \(G\), define

\[
\ell_{\rho \alpha}[g] := \inf_{\phi[C]=g} \ell_{\rho \alpha}[C]
\]

\[
\text{EL}'_{\alpha}[g] := \sup_{\rho: \text{Edge}(\Gamma) \rightarrow \mathbb{R}_{\geq 0}} \frac{\ell_{\rho \alpha}[g]^2}{\text{Area}_{\rho \alpha}(\Gamma)}
\]

\[
\text{EL}_{\alpha}[g] := \lim_{n \rightarrow \infty} \text{EL}'_{\alpha}[g^n]/n^2.
\]

There are also other models for defining a combinatorial length of curves on a graph. Notably, Schramm [Sch93] and Canon, Floyd, and Parry [CFP94] define a model where the length of a curve in a graph is determined by the vertices that it passes through, rather than the edges that it crosses over (as in this paper). More generally, one can consider *shingleings* of a graph or surface, decompositions of the space into a finite number of overlapping open sets. The combinatorial length of a curve is then determined by which shingles it passes through. The edge model that is the main focus of this paper comes from taking shingles that are neighborhoods of the edges (overlapping at the vertices), while the vertex model of Schramm–Cannon–Floyd–Perry comes from taking shingles that are neighborhoods of the vertices (overlapping at the centers of edges).

For any of these notions of length, one can define a notion of extremal length using Equation (5.10).

More generally, one may instead attempt to characterize which extremal length functions can appear. There are several natural sources of an “extremal length” function on \( \mathcal{M} \mathcal{F}^+(\Sigma) \):

1. A conformal structure \( \omega \) on \( \Sigma \) gives the usual notion of extremal length.
2. If \( \Sigma' \) is another surface and \( \phi: \Sigma' \hookrightarrow \Sigma \) is a filling embedding (an embedding for which the complementary regions are disks or annuli), then a conformal structure \( \omega' \) on \( \Sigma' \) gives a notion of extremal length on \( \Sigma \), defined analogously to Definition 5.23.
3. An elastic graph spine \((\Gamma, \alpha)\) for \( \Sigma \) gives a notion of extremal length by Equations (5.10) or (5.13).
4. A filling elastic graph in \( \Sigma \) gives a notion of extremal length by Definition 5.23.
5. Finally, a shingling of \( \Sigma \) gives yet another notion of extremal length.

All of these notions of extremal length give a function on \( \mathcal{C}^+(\Sigma) \) that is

- positive,
- homogeneous quadratic, in the sense of Lemma 5.4,
- does not increase under smoothing, in the sense of Lemma 5.5,
- sub-additive under union, in the sense of Lemma 5.6, and therefore
- extends to a convex function on \( \mathcal{M} \mathcal{F}^+(\Sigma) \), by Theorem 2.

As a result of the convexity, we can think of EL as a kind of norm on \( \mathcal{M} \mathcal{F}^+(\Sigma) \).
Problem 5.25. Which functions \( \text{EL} : \mathcal{MF}^+(\Sigma) \to \mathbb{R} \) satisfying the properties above can arise from the constructions above?

Some of these notions of extremal length subsume the others: Extremal lengths from notion (5)) are dense extremal lengths from notion (4) by a direct construction. By Proposition 5.15, notion (2) is dense in notion (4) (up to scale). Notion (4) naturally includes notion (3), and by taking the graph to be the edges of a triangulation we can see that notion (4) is dense in notions (1) and (2).

Problem 5.26. The definition of extremal length in surfaces, Equation (5.1), extends to (width) graphs embedded in \( \Sigma \), rather than just curves. What do the resulting optimal metrics look like?

6. Stretch factors and embedding energy

6.1. Stretch factors for surfaces. Let \( \phi : \Sigma_1 \to \Sigma_2 \) be a topological embedding of surfaces. Then composing with \( \phi \) and deleting null-homotopic components induces a natural map \( \phi_* : C^+(\Sigma_1) \to C^+(\Sigma_2) \). (This pushforward does not work on \( C^-(\Sigma_1) \).)

Warning 6.1. The map \( \phi_* \) does not extend to a continuous map \( \mathcal{MF}^+(\Sigma_1) \to \mathcal{MF}^+(\Sigma_2) \).

Now suppose \( \Sigma_1 \) and \( \Sigma_2 \) have conformal structures \( \omega_1 \) and \( \omega_2 \), respectively. (The map \( \phi \) need not respect the conformal structures.)

Definition 6.2. In the above setting, the stretch factor of \( \phi \) is

\[
(6.3) \quad \text{SF}[\phi] := \sup_{C \in C^+(\Sigma)} \frac{\text{EL}_{\omega_2}[\phi_* C]}{\text{EL}_{\omega_1}[C]}. 
\]

This depends only on the homotopy class of \( \phi \).

It follows from the definition that SF behaves well under composition.

Proposition 6.4. If \( f : \Sigma_1 \hookrightarrow \Sigma_2 \) and \( g : \Sigma_2 \hookrightarrow \Sigma_3 \) are two topological embeddings of conformal surfaces, then

\[ \text{SF}[f \circ g] \leq \text{SF}[f] \cdot \text{SF}[g]. \]

Definition 6.5. A conformal embedding \( \phi : \Sigma_1 \hookrightarrow \Sigma_2 \) is strict if, in each component of \( \Sigma_2 \), there is a non-empty open subset in the complement of the image.

Theorem 3 (essentially Ioffe). If \( (\Sigma_1, \omega_1) \) and \( (\Sigma_2, \omega_2) \) are two conformal surfaces and \( \phi : \Sigma_1 \to \Sigma_2 \) is a homeomorphism so that there is no strict conformal embedding in the homotopy class \( [\phi] \), let \( K \) be the lowest constant so that there is a \( K \)-quasi-conformal map in the homotopy class \( [\phi] \). Then

\[ \text{SF}[\phi] = K. \]

Proof. This is very close to the main theorem of [Iof75]. In that paper, Ioffe proves that if there is no conformal embedding in \( [\phi] \), there are canonical quadratic differentials \( q_i \in Q^+(\Sigma_i) \) and a representative for \( \phi \) that uniformly stretches \( q_1 \) to \( q_2 \). Suitably approximating the horizontal foliations of \( Q_1 \) and \( Q_2 \) by rational measured foliations (being careful about Warning 6.1) gives a sequence of simple curves (not necessarily connected) so the ratio of extremal lengths approaches \( K \).
The case when there is a conformal embedding, but not a strict conformal embedding, can be treated by, for instance, adding annuli to the boundary components of $\Sigma_1$ so there is no conformal embedding. □

Theorem 3 shows, in particular, that the stretch factor generalizes the usual Teichmüller metric: if $\Sigma_1$ and $\Sigma_2$ are closed surfaces, then there is never a conformal embedding of $\Sigma_1$ into $\Sigma_2$ unless they are equal, and $\log \text{SF}[\phi]$ is Teichmüller distance between $\Sigma_1$ and $\Sigma_2$. In this context, Proposition 6.4 is the triangle inequality for Teichmüller distance. Theorem 3 for closed surfaces was also proved by Kerckhoff [Ker80, Theorem 4]. In this case connected simple curves suffice.

Remark 6.6. Despite Warning 6.1, there is a non-continuous extension of $\phi_*$ to measured foliations. Define a pull-back function $\phi^*: \mathcal{C}^-(\Sigma_2) \to \mathcal{M}\mathcal{F}^-(\Sigma_1)$ by taking all possible intersections of $C$ with the image of $\phi_*(\Sigma_1)$:

$$\phi^*(C) := \{ [c \cup \phi(\Sigma_1)] \mid c \text{ a representative of } C \},$$

where we delete inadmissible components as in the definition of $\phi_*$. Then, for a measured foliation $F_1 \in \mathcal{M}\mathcal{F}^+(\Sigma_1)$, we may define $\phi_*(F_2)$ to be the (unique) measured foliation $F_2 \in \mathcal{M}\mathcal{F}^+(\Sigma_2)$ so that, for all $C_2 \in \mathcal{C}^-(\Sigma_2)$,

$$i([C_2], F_2) = \inf_{C_1 \in \phi^*(C_2)} i([C_1], F_1).$$

With this definition, the curves in Equation (5.1) can be replaced by measured foliations, and the supremum is achieved.

Definition 6.7. The embedding $\phi$ is annular if it extends to an embedding of an annular extension $\hat{\Sigma}_1$ in $\Sigma_2$, where $\hat{\Sigma}_1$ is obtained by attaching an annulus to each boundary component of $\Sigma_1$.

Theorem 4 (Joint with K. Pilgrim). If $(\Sigma_1, \omega_1)$ and $(\Sigma_2, \omega_2)$ are two conformal surfaces and $\phi: \Sigma_1 \to \Sigma_2$ is a topological embedding, then $\phi$ is homotopic to a conformal embedding iff $\text{SF}[\phi] \leq 1$.

Furthermore, the following conditions are equivalent:

(1) $\text{SF}[\phi] < 1$,

(2) $\phi$ is homotopic to an annular conformal embedding,

(3) $\phi$ is homotopic to a strict conformal embedding, and

Proof sketch. The case when $\phi$ is not homotopic to an embedding is implied by Theorem 3. The converse of the first claim is essentially Schwarz’s Lemma.

To show that (1) implies (2), pick a quadratic differential $q \in Q^+(\Sigma_1)$ that is strictly positive on each boundary component. Define $\hat{\Sigma}_1'$ to be $\Sigma_1$ plus an annulus of width $t$ on each boundary component, using the coordinates from $q$. Elementary estimates show that $\text{SF}[\hat{\Sigma}_1' \to \Sigma_1]$ (in the natural homotopy class) approaches 1 as $t$ approaches 0. Then for $t$ small, by Proposition 6.4

$$\text{SF}[\hat{\Sigma}_1' \to \Sigma_2] \leq \text{SF}[\hat{\Sigma}_1' \to \Sigma_1] \cdot \text{SF}[\Sigma_1 \to \Sigma_2] < 1,$$

so by Theorem 3 there is a conformal embedding of $\hat{\Sigma}_1'$ in $\Sigma_2$. 
Clearly (2) implies (3). Finally, if \( \phi \) is a strict conformal embedding of \( \Sigma_1 \) in \( \Sigma_2 \), then by considering test metrics we can show that

\[
SF[\phi] \leq \sup_{q \in \mathbb{Q}^+(\Sigma_2)} \frac{\text{Area}_q(\text{Im}(\phi))}{\text{Area}_q(\Sigma_2)}.
\]

Here \( \text{Im}(\phi) \subset \Sigma_2 \) is the image of \( \phi \), which by hypothesis misses an open subset of \( \Sigma_2 \), and \( \text{Area}_q \) is the area with respect to the quadratic differential \( q \). Thus for each \( q \), \( \text{Area}_q(\text{Im}(\phi))/\text{Area}_q(\Sigma_2) < 1 \). Since the suprema doesn’t change as we scale \( q \), we are maximizing over the compact set \( \mathbb{P}\mathbb{Q}^+(\Sigma_2, \omega_2) \) and the supremum is strictly less than 1.

6.2. **Behaviour of stretch factor under covers.** In Section 7.2, we will also need an understanding of the behavior of SF under covers. Let \( \phi : \Sigma_1 \hookrightarrow \Sigma_2 \) be a topological embedding of conformal surfaces, let \( \tilde{\Sigma}_2 \) be a finite cover of \( \Sigma_2 \), and let \( \tilde{\phi} : \tilde{\Sigma}_1 \to \tilde{\Sigma}_2 \) be the pull-back of \( \phi \); that is, \( \tilde{\Sigma}_1 \) is defined by the pull-back of \( \tilde{\Sigma}_2 \) and \( \Sigma_1 \) via the diagram

\[
\begin{array}{ccc}
\tilde{\Sigma}_1 & \xrightarrow{\tilde{\phi}} & \tilde{\Sigma}_2 \\
\downarrow & & \downarrow \\
\Sigma_1 & \xrightarrow{\phi} & \Sigma_2.
\end{array}
\]

When \( \tilde{\Sigma}_2 \) is connected, \( \tilde{\Sigma}_1 \) is the minimal cover of \( \Sigma_1 \) with a map to \( \tilde{\Sigma}_2 \) making the above diagram commute.

**Question 6.10.** How does SF[\( \tilde{\phi} \)] compare to SF[\( \phi \)]? It appears that SF[\( \tilde{\phi} \)] \( \neq \) SF[\( \phi \)] in general, at least if we allow \( \Sigma_2 \) to be an orbifold. However, we can make many partial statements.

**Lemma 6.11.** For \( \phi \) and \( \tilde{\phi} \) as above, SF[\( \tilde{\phi} \)] \( \geq \) SF[\( \phi \)].

**Proof.** Follows from the definition of SF and the good behavior of extremal length under covers, Lemma 5.3. \( \square \)

**Lemma 6.12.** For \( \phi \) and \( \tilde{\phi} \) as above, if SF[\( \phi \)] \( \geq 1 \), then SF[\( \tilde{\phi} \)] = SF[\( \phi \)].

**Proof.** Follows from Theorem 3. \( \square \)

**Lemma 6.13.** For \( \phi \) and \( \tilde{\phi} \) as above, SF[\( \tilde{\phi} \)] \( < 1 \) iff SF[\( \phi \)] \( < 1 \) and SF[\( \tilde{\phi} \)] = 1 iff SF[\( \phi \)] = 1.

**Proof.** If SF[\( \phi \)] \( < 1 \), by Theorem 4 the map \( \phi \) is homotopic to a strict conformal embedding, and so \( \tilde{\phi} \) is too and thus SF[\( \tilde{\phi} \)] \( < 1 \). The rest follows from Lemma 6.12. \( \square \)

**Definition 6.14.** For \( \phi : \Sigma_1 \to \Sigma_2 \) a topological embedding of surfaces, define the **lifted stretch factor** by

\[
\{SF[\phi] := \lim_{\tilde{\phi} \text{ covers } \phi} SF[\tilde{\phi}].
\]
where the limit runs over increasingly large covers of \( \phi \) (determined by a covering of \( \Sigma_2 \)). These covers form a directed system, and SF only increases in a cover (Lemma 6.11) while remaining bounded (Lemmas 6.12 and 6.13), so the limit exists.

By definition, if \( \tilde{\phi} \) is a covering of \( \phi \), then \( \tilde{\text{SF}}[\tilde{\phi}] = \tilde{\text{SF}}[\phi] \).

We will ultimately need to show that certain maps have lifted stretch factor less than one. We give two arguments, one more elementary and one proving a stronger result. (See the two proofs of Proposition 7.10.)

**Definition 6.15.** A homotopy class of topological embeddings \([\phi]: \Sigma_1 \to \Sigma_2\) is **conformally loose** if, for all \( y \in \Sigma_2 \), there is a conformal embedding \( \psi \in [\phi] \) so that \( y \notin \phi(\Sigma_1) \). If \( \Sigma_2 \) is compact and \([\phi]\) is conformally loose, then we can find finitely many conformal embeddings \( \phi_i \in [\phi], i = 1, \ldots, n \) so that

\[
\bigcap_{i=1}^{n} \phi(\Sigma_1) = \emptyset.
\]

**Proposition 6.17.** If \([\phi]\) is conformally loose with \( n \) maps \( \phi_i \) in Equation (6.16), then \( \text{SF}[\phi] \leq 1 - 1/n \).

*Proof sketch.* Fix a weighted multi-curve \( C \) on \( \Sigma_1 \), and let \( q \in Q^+(\Sigma_2, \omega_2) \) be the quadratic differential corresponding to \( \phi(C) \). For at least one \( i \), we will have

\[
\text{Area}_q(\phi_i(\Sigma_1)) \leq 1 - 1/n.
\]

Choosing test metrics as in Equation (6.8) shows that \( \text{EL}_{\omega_1}[C] \leq (1-1/n) \text{EL}_{\omega_2}[\phi_\ast C] \), as desired.

**Corollary 6.18.** If \([\phi]\) is conformally loose, then \( \tilde{\text{SF}}[\phi] < 1 \).

Corollary 6.18 is enough to prove Theorem 1. However, more is true.

**Theorem 5** (Joint with J. Kahn and K. Pilgrim). *For every strict conformal embedding \( \phi: \Sigma_1 \hookrightarrow \Sigma_2 \),

\[
\tilde{\text{SF}}[\phi] < 1.
\]

By Equation (6.8), Theorem 5 is a consequence of the following proposition.

**Proposition 6.19.** Let \( \Sigma_2 \) be a Riemann surface, and let \( \Sigma_1 \subset \Sigma_2 \) be a subsurface with compact closure. Then there is a constant \( K < 1 \) so that, for all \( 0 \neq q \in Q(\Sigma_2) \),

\[
\frac{\text{Area}_q(\Sigma_1)}{\text{Area}_q(\Sigma_2)} < K.
\]

Furthermore, \( K \) can be chosen so that Equation (6.20) holds for all coverings of the pair \((\Sigma_1, \Sigma_2)\).

Proposition 6.19, in turn, depends on the following local lemma.

**Lemma 6.21.** Let \( \Omega \subset \mathbb{D} \) be an open subset of the disk with an open set \( D \) in the complement of \( \Omega \), and let \( B \subset \mathbb{D} \) be a neighborhood of \( \overline{\Omega} \cap \partial \mathbb{D} \). Then, for every \( \epsilon > 0 \), there is a \( \delta > 0 \) so that, if \( q \in Q(\mathbb{D}) \) is such that

\[
\frac{\text{Area}_q(\Omega)}{\text{Area}_q(\mathbb{D})} > 1 - \delta,
\]
then

\[
\frac{\text{Area}_q(B)}{\text{Area}_q(\mathbb{D})} > 1 - \epsilon.
\]

Essentially, Lemma 6.21 says that if the area of \( q \) is concentrating in a subset of the disk, then it is concentrating near the boundary. See Figure 5.

**Proof sketch for Lemma 6.21.** If there are no such bounds, there is an \( \epsilon \) so that we can find a sequence of quadratic differentials \( q_n \in \mathcal{Q}(\mathbb{D}) \) so that

\[
\text{Area}_{q_n}(\mathbb{D}) = 1
\]

\[
\text{Area}_{q_n}(B) < \epsilon
\]

\[
\text{Area}_{q_n}(\Omega) > 1 - 1/n.
\]

Consider \( |q_n| \) as a measure on \( \mathbb{D} \). Since the space of measures of unit area on the closed disk is compact in the weak topology, after passing to a subsequence we may assume that \( |q_n| \) converges in the weak topology to some limiting measure \( \mu \) on \( \mathbb{D} \). Absolute values of holomorphic functions are closed in the weak topology, so the restriction of \( \mu \) to the open disk can be written \( |q_x| \) for some holomorphic quadratic differential \( q_x \). But \( \text{Area}_{q_x}(D) < 1/n \), so \( \text{Area}_{q_x}(D) = 0 \), so \( q_x \) is identically 0; hence \( \mu \) is supported on \( \partial \mathbb{D} \). Equation (6.26) implies that the support of \( \mu \) is also contained in \( \overline{\Omega} \), and hence in \( \overline{\Omega} \cap \partial \mathbb{D} \). But this contradicts Equation (6.25).

**Proof sketch for Proposition 6.19.** Divide \( \Sigma_2 \) by smooth arcs \( \alpha_j \) so that the complementary regions are all disks; let these disks be \( U_i \). Consider a quadratic differential on \( \Sigma_2 \) with a very large proportion of its area in \( \Sigma_1 \). Then most \( U_i \) (as weighted by \( \text{Area}_q(U_i) \)) must have most of their \( q \)-area in \( U_i \cap \Sigma_1 \). Lemma 6.21 then says that most of the area of most of the \( U_i \) must be in a small neighborhood of the seams \( \alpha_j \). Arrange the constants so that more than half the total area must be in these neighborhoods.

Now pick an alternate set of seams \( \beta_k \) with the same properties, but so that they never overlap. By the same argument, more than half the total area must be concentrated in a neighborhood of the \( \beta_k \), a contradiction.

All of the estimates in this proof only depend on the local geometry of the \( U_i \), and thus remain unchanged under taking covers.

**6.3. Stretch factors and embedding energy for graphs.** We now turn to the (easier) parallel theory of stretch factors for maps between graphs. Let \( \phi: \Gamma_1 \to \Gamma_2 \)
be a continuous map between two graphs, and suppose that \( \Gamma_1 \) and \( \Gamma_2 \) have elastic structures \( \alpha_1 \) and \( \alpha_2 \), respectively.

**Definition 6.27.** In this setting, the EL stretch factor of \( \phi \) is

\[
SF_{\text{EL}}[\phi] := \sup_{C \text{ curve on } \Gamma_1} \frac{\text{EL}_{\alpha_2}[\phi_C C]}{\text{EL}_{\alpha_1}[C].}
\]

As for surfaces, this only depends on the homotopy class of \( \phi \).

Dually, the Dirichlet stretch factor of \( \phi \) is

\[
SF_{\text{Dir}}[\phi] := \sup_{K \text{ length graph}} \frac{\text{Dir}[\phi \circ f]}{\text{Dir}[f]},
\]

where the supremum runs over all length graphs \((K, \ell)\) and all homotopy classes of maps from \( \Gamma_2 \) to \( K \).

**Remark 6.30.** As mentioned in the introduction, the Dirichlet stretch factor has a natural interpretation in terms of rubber-band networks. If \( SF_{\text{Dir}}[\phi] < 1 \), then the rubber-band network \( \Gamma_1 \) is "looser" than the rubber-band network \( \Gamma_2 \), in the sense that however \( \Gamma_2 \) is stretched over any target graph \( K \) and any map \( f \) from \( \Gamma_2 \) to \( K \), the minimum Dirichlet energy of \( \Gamma_1 \) in the homotopy class \([\phi \circ f]\) is less than the minimum Dirichlet energy of \( \Gamma_2 \) in the homotopy class \( [f] \). (Theorem 6 below implies that the same is true for arbitrary target metric spaces, not just graphs.)

As in the case of surfaces, these stretch factors behave well under composition (Proposition 6.4). Unlike in the case of surfaces, for graphs we have a direct characterization of the stretch factor.

**Definition 6.31.** For \( \phi : \Gamma_1 \rightarrow \Gamma_2 \) a Lipshitz map between elastic graphs, the embedding energy of \( \phi \) is

\[
\text{Emb}(\phi) := \text{ess sup}_{y \in \Gamma_2} \text{Fill}_\phi(y)
\]

\[
\text{Emb}[\phi] := \inf_{\psi \in [\phi]} \text{Emb}(\psi),
\]

where \( \text{Fill}_\phi \) is the filling function defined in Equation (4.5). We take the \( L^\infty \) norm of \( \text{Fill}_\phi \) rather than the \( L^1 \) norm used for Dirichlet energy. This is equivalent to Equation (1.2).

By comparing Equations (1.2) and (4.11), we see that

\[
\text{Emb}(\phi) \geq \text{Lip}(\phi).
\]

**Theorem 6.** For \( \phi : \Gamma_1 \rightarrow \Gamma_2 \) a continuous map between elastic graphs,

\[
SF_{\text{EL}}[\phi] = SF_{\text{Dir}}[\phi] = \text{Emb}[\phi].
\]

Theorem 6 should be thought of as analogous to Theorem 3, although it applies in all cases, not just when there fails to be a conformal embedding. It is also analogous to the relation between Lipshitz energy as maximum derivative (Equation (4.11)) and as ratio of curve lengths (Equation (4.13)).
Proof sketch. For $C$ a curve on $\Gamma_1$ and $\phi: \Gamma_1 \to \Gamma_2$ a map, by pushing forward the scaling function $\rho$ in Equation (5.10) we can see that

\begin{equation}
(6.35) \quad \text{EL}[\phi_* C] \leq \text{EL}[C] \cdot \text{Emb}(\phi).
\end{equation}

This immediately implies that $\text{Emb}[\phi] \leq \text{SF}_{\text{EL}}[\phi]$.

Similarly, the definition of the Dirichlet energy as the $L^1$ norm of the filling function and the embedding energy as the $L^\infty$ norm of the filling function makes it clear that, for $f: \Gamma_2 \to K$ any map to a length graph,

\begin{equation}
(6.36) \quad \text{Dir}(f \circ \phi) \leq \text{Dir}(f) \cdot \text{Emb}(\phi),
\end{equation}

which implies that $\text{Emb}[\phi] \leq \text{SF}_{\text{Dir}}[\phi]$.

To prove the opposite inequalities, it suffices to find a representative for the homotopy class $[\phi]$ for which Equations (6.35) and (6.36) are tight. We defer the description of these nice representatives (analogous to harmonic maps, Definition 4.6) until Section 6.5. □

In contrast with the case for surfaces (Section 6.2), it is easy to see how the stretch factor for graphs behaves under covers.

**Proposition 6.37.** Let $\phi: \Gamma_1 \to \Gamma_2$ be a map of elastic graphs, $\tilde{\Gamma}_2$ be a cover of $\Gamma_2$, and $\tilde{\phi}: \tilde{\Gamma}_1 \to \tilde{\Gamma}_2$ be the pull-back map, defined as in Equation (6.9). Then

$$\text{SF}[\tilde{\phi}] = \text{SF}[\phi].$$

**Proof.** The definition of $\text{SF}[\phi]$ by the maximal expansion of extremal length or Dirichlet energy shows that $\text{SF}[\tilde{\phi}] \geq \text{SF}[\phi]$ (as in Lemma 6.11). The definition of $\text{Emb}[\phi]$ by the infimum over all representatives shows that $\text{Emb}[\tilde{\phi}] \leq \text{Emb}[\phi]$. Theorem 6 shows that we have equalities. □

6.4. **Relating graphs and surfaces.** As motivation for the somewhat strange definition of embedding energy, consider two elastic graphs $(\Gamma_1, \alpha_1)$ and $(\Gamma_2, \alpha_2)$ and a conformal embedding of the thickenings $N_t \Gamma_1 \hookrightarrow N_t \Gamma_2$. Suppose that this conformal embedding is “close” in some sense to a graph map $\phi: \Gamma_1 \to \Gamma_2$. If we look away from the thickenings of the vertices (of both graphs), we see locally a map from some edges of $\Gamma_1$ into a single edge of $\Gamma_2$. The total width of images of the thickened edges of $\Gamma_1$ must be less than or equal to the total available width in the thickened edge of $\Gamma_2$. Near a point $x \in \Gamma_1$, the image is stretched horizontally by a factor of $|\phi'(x)|$; since we are considering a conformal embedding, the image must also be stretched vertically by the same factor. Thus the image of this portion of the edge takes up a width of $t|\phi'(x)|$. This argument suggests that if there is a conformal embedding close to $\phi$, we must have, for each $y \in \Gamma_2$ not near a vertex or the image of a vertex,

$$\sum_{\phi(x) = y} t|\phi'(x)| \leq t \quad \text{or, equivalently,} \quad \text{Fill}_\phi(y) \leq 1.$$

See Figure 6.

We will not attempt to make the above heuristic argument precise. Instead, we get a precise statement another way.
Figure 6. Motivation for the definition of embedding energy. Two edges of \( \Gamma_1 \) mapping to one edge of \( \Gamma_2 \) get thickened up to two rectangles mapping to one rectangle. For the result to be a conformal embedding, we need the total width of the image to be less than the total width available in the range.

**Proposition 6.38.** Let \( \Gamma_1 \) and \( \Gamma_2 \) be trivalent elastic ribbon graphs, and let \( m = \min_e \alpha(e) \) be the smallest weight of any edge in either \( \Gamma_1 \) or \( \Gamma_2 \). Let \( \phi: \Gamma_1 \to \Gamma_2 \) be a map that extends to a topological embedding \( N_t\phi: N_t\Gamma_1 \to N_t\Gamma_2 \). Then, for \( t < m/2 \), we have

\[
SF[\phi]/(1 + 8t/m) \leq SF[N_t\phi] \leq SF[\phi] \cdot (1 + 8t/m).
\]

**Proof.** Immediate from Proposition 5.15. \( \square \)

**Theorem 7.** Let \( \Gamma_1 \) and \( \Gamma_2 \) be elastic ribbon graphs, and let \( \phi: \Gamma_1 \to \Gamma_2 \) be a map of graphs that extends to a topological embedding \( N_\phi: N\Gamma_1 \to N\Gamma_2 \). If \( \text{Emb}[\phi] < 1 \), then for all \( t \) sufficiently small, \( N_t\Gamma_1 \) conformally embeds in \( N_t\Gamma_2 \) in \( [N_\phi] \). Conversely, if \( N_t\Gamma_1 \) conformally embeds in \( N_t\Gamma_2 \) in \( [N_\phi] \) for all sufficiently small \( t \), then \( \text{Emb}[\phi] \leq 1 \).

**Proof.** Follows from Proposition 6.38 and Theorem 4. If the graphs \( \Gamma_1 \) and \( \Gamma_2 \) are not trivalent, we may approximate them by trivalent graphs at the cost of a small change in the stretch factor. \( \square \)

Motivated by Theorem 7, we make the following definition.

**Definition 6.39.** For \( (\Gamma_1, \alpha_1) \) and \( (\Gamma_2, \alpha_2) \) two elastic strip graphs, a map \( \phi: \Gamma_1 \to \Gamma_2 \) is a **conformal embedding** if \( \text{Emb}[\phi] \leq 1 \). The conformal embedding is **strict** if \( \text{Emb}[\phi] < 1 \). We say that \( \Gamma_1 \) (strictly) **conformally embeds** in \( \Gamma_2 \) in the homotopy class \([\phi]\) if there is a (strict) conformal embedding in \([\phi]\).

A conformal embedding of graphs is neither an embedding of graphs, nor conformal in any obvious sense.

6.5. **Minimizing embedding energy: \( \lambda \)-filling maps.** We say a little more about the proof of Theorem 6, and in particular characterize the optimal maps. Recall from Section 3.3 that a strip graph has both lengths and widths.

**Definition 6.40.** For \( \lambda > 0 \), a map \( \phi: S_1 \to S_2 \) between strip graphs is **\( \lambda \)-filling** if
(1) \( \phi \) is length-preserving: it does not backtrack, and \( |\phi'(x)| = 1 \) for almost all \( x \), where we take derivative with respect to the length metrics, and
(2) \( \phi \) scales widths by \( \lambda \): for all \( y \in S_2 \),
\[
\sum_{\phi(x) = y} w(x) = \lambda \cdot w(y).
\]

Note that for maps between strip graphs, we have to be careful whether we differentiate with respect to the length metric or with respect to the coordinates from the elastic weights, as in Warning 3.11.

**Lemma 6.41.** If \( \phi: S_1 \to S_2 \) is a \( \lambda \)-filling map between strip graphs, then the underlying map \( \Gamma(\phi): \Gamma(S_1) \to \Gamma(S_2) \) between elastic graphs has embedding energy \( \lambda \). This is minimal in the homotopy class.

**Proof sketch.** The conditions on \( \phi \) imply that \( \text{Fill}_{\Gamma(\phi)} = \lambda \) everywhere on \( \Gamma(S_2) \). (Recall that “length” on the elastic graph is length/width in terms of the strip graph.)

Now consider the length graph \( K(S_2) \). The map \( \Gamma(S_2) \to K(S_2) \) has Dirichlet energy \( \text{Area}(S_2) \), almost by definition. The composite map \( \Gamma(S_1) \to \Gamma(S_2) \to K(S_2) \) has Dirichlet energy \( \lambda \cdot \text{Area}(S_2) \). It follows that \( \text{SF}_{\text{Dir}}[\Gamma(\phi)] \geq \lambda \). The easy direction of Theorem 6 then implies that \( \text{Emb}[\Gamma(\phi)] = \lambda \).

**Proposition 6.42.** Let \( \phi: \Gamma_1 \to \Gamma_2 \) be a map between elastic graphs. Then there are strip structures \( S_1 \) on \( \Gamma_1 \) and \( S_2 \) on \( \Gamma_2 \) so that \( \phi \) is \( \lambda \)-filling as a map between strip graphs iff the following conditions are satisfied.

(1) \( \text{Fill}_{\phi}(y) = \lambda \) everywhere on \( \Gamma_2 \).
(2) If \( x_1 \) and \( x_2 \) are in the same edge of \( \Gamma_1 \) and \( \phi(x_1) \) and \( \phi(x_2) \) are in the same edge of \( \Gamma_2 \), then
\[
|\phi'(x_1)| = |\phi'(x_2)|.
\]
(3) If the pairs \( (x_1, y_1) \) and \( (x_2, y_2) \) are each on a single edge of \( \Gamma_1 \) with \( \phi(x_1) = \phi(y_1) \) and \( \phi(x_2) = \phi(y_2) \), then
\[
\frac{|\phi'(x_1)|}{|\phi'(x_2)|} = \frac{|\phi'(y_1)|}{|\phi'(y_2)|}.
\]
(4) Near a vertex \( v \) of \( \Gamma_1 \), the quantities \( |\phi'(x)| \) satisfy the mapped triangle inequalities of Conditions (4) and (5) of Definition 4.6.

In view of Proposition 6.42, we say that a map between elastic graphs is \( \lambda \)-filling if it satisfies the condition there. It is not true that every homotopy class of maps between elastic graphs has a \( \lambda \)-filling representative. However, we can make it \( \lambda \)-filling on a subgraph.

**Definition 6.43.** A map \( \phi: S_1 \to S_2 \) between strip surfaces is partially \( \lambda \)-filling if there are non-empty subgraphs \( T_1 \) of \( S_1 \) and \( T_2 \) of \( S_2 \) so that
(1) \( \phi(T_1) = T_2 \) and \( \phi^{-1}(T_2) = T_1 \);
(2) the restriction of \( \phi \) to a map \( T_1 \to T_2 \) is \( \lambda \)-filling;
(3) \( \phi \) is everywhere length-preserving; and
(4) outside of \( T_1 \) and \( T_2 \), the map \( \phi \) scales widths by less than \( \lambda \).

**Proposition 6.44.** Every homotopy class of maps between elastic graphs has a partially \( \lambda \)-filling representative.
Theorem 6 follows quickly from Proposition 6.44.

6.6. Computing embedding energy. Given two elastic graphs \( \Gamma_1 \) and \( \Gamma_2 \) and a homotopy class \([\phi]: \Gamma_1 \to \Gamma_2\) of maps between them, how can we concretely find the partially \( \lambda \)-filling representative \( \psi \in [\phi] \) guaranteed by Proposition 6.44? The following iteration appears to converge.

(0) Pick a set of widths \( v_0 \in \mathcal{W}(\Gamma_2) \), i.e., a width for each edge of \( \Gamma_2 \) satisfying the triangle inequality.

(1) Take the metric graph \( K_0 \) to be \( \Gamma_2 \) with edge \( e \) assigned length \( \alpha(e)v_0(e) \). Note that the evident map \( f_0: \Gamma_2 \to K_0 \) is harmonic.

(2) Find a harmonic representative \( g_0 \) of the composite map \( \Gamma_1 \xrightarrow{\phi} \Gamma_2 \xrightarrow{f} K_0 \). Our first approximation to \( \psi \) is \( \psi_0 = f_0 \circ g_0 \).

(3) Compute the tension of \( g_0 \) (i.e., \( |g_0'(x)| \)) in each edge of \( \Gamma_1 \), getting \( w_0 \in \mathcal{W}(\Gamma_1) \).

(4) Push forward \( w_0 \) to a function \( v_1 \in \mathcal{W}(\Gamma_2) \) by setting, for \( e \in \text{Edge}(\Gamma_2) \) and \( y \in e \),

\[
v_1(y) = \sum_{x \in g^{-1}(f(y))} w_0(x).
\]

This is independent of the choice of \( y \) since \( g \) is harmonic. Then return to Step (1), using \( v_1 \) instead of \( v_0 \).

Schematically, we are iterating around the following cycle:

\[
\begin{array}{ccc}
\mathcal{L}(\Gamma_2) & \stackrel{(1)}{\to} & \mathcal{W}(\Gamma_2) \\
\downarrow & & \downarrow \\
\mathcal{L}(\Gamma_1) & \stackrel{(3)}{\to} & \mathcal{W}(\Gamma_1) \\
\end{array}
\]

At each step, we can compute the embedding energy:

\[
\text{Emb}(\psi_i) = \max_{e \in \text{Edge}(\Gamma_i)} \frac{v_{i+1}(e)}{v_i(e)}.
\]

**Conjecture 6.45.** The algorithm above converges to a map with lowest embedding energy.

Once the combinatorics of the graph map have settled down, this maps in this iteration become linear, and the algorithm reduces to finding the largest eigenvector of a linear system by iteration. In practice, the algorithm appears to converge rapidly.

Note the relation to tightness in Equation (1.13): we simultaneously find a representative \( \psi \in [\phi] \), a map \( f \) from \( \Gamma_2 \) to a metric graph \( K \), and a set of widths \( w \) on \( \Gamma_1 \), all multiplicative on the nose:

\[
\text{Emb}(\psi) = \frac{\text{Dir}[f \circ \psi]}{\text{Dir}[f]} = \frac{\text{EL}[\psi_*w]}{\text{EL}[w]}.
\]

6.7. Extensions and questions.

**Conjecture 6.46.** For any strict conformal embedding \( \phi: \Sigma_1 \to \Sigma_2 \), there is a cover \( \tilde{\phi}: \tilde{\Sigma}_1 \to \tilde{\Sigma}_2 \) so that \([\tilde{\phi}]\) is conformally loose.
**Question 6.47.** What happens if we vary the definition of SF for surfaces? For instance, we restricted to simple curves in Definition 6.2. What happens if we drop that restriction, and look at general curves? What if we look at the expansion factor of Dirichlet energy for maps to graphs instead? What if we look at width graphs, as in Problem 5.26?

**Problem 6.48.** Give a direct expression for (some version of) the stretch factor SF[φ] or lifted stretch factor SF[φ] of a conformal surface embedding, analogous to Theorem 3 for the case when φ is not homotopic to a conformal embedding or to Theorem 6 for case of graphs.

**Problem 6.49.** In Definition 5.23, we defined a notion of extremal length starting from an elastic graph Γ embedding as a filling subset of a surface. Give a direct expression for the stretch factor (maximal ratio of extremal lengths) between two graphs Γ₁ and Γ₂ with filling embeddings in the same surface Σ. (Theorem 6 handles the case when Γ₂ is a spine.)

Since filling graphs can be used to approximate surfaces, a solution to Problem 6.49 would presumably be helpful in answering Problem 6.48.

**Problem 6.50.** Extend Problem 6.49 to the more general setting of a group G and surjective maps π₁ : Γ₁ → G and π₂ : Γ₂ → G, as in Definition 5.24.

7. **Dynamics**

7.1. **Iterating covers.** We finally turn to the dynamical picture: what happens when we iterate a map from a conformal surface or elastic graph to itself? Our setting is not quite the usual dynamical picture: we are “iterating” virtual endomorphisms, which are not maps from a space X to itself, but maps from a cover ̃X to X.

**Definition 7.1.** Let X₀ and X₁ be topological spaces, π : X₁ → X₀ be a covering map of degree d, and φ : X₁ → X₀ be a continuous map. We call this data a virtual endomorphism of X₀, also called a topological automaton by Nekrashevych [Nek05, Nek14]. Define Xₖ to be the k-fold product of X₁ with itself over X₀ using the two maps π and φ; e.g., X₃ is the pullback of the diagram below.

```
\[
\begin{array}{c}
X_3 \\
\pi_3 \\
\end{array}
\downarrow
\begin{array}{c}
X_1 \\
\phi \\
\end{array}
\downarrow
\begin{array}{c}
X_0 \\
\pi \\
\end{array}
\]
```

Concretely, define

\[X_k := \{ (x_1, \ldots, x_k) \in (X_1)^k \mid \phi(x_i) = \pi(x_{i+1}) \}\]

Xₖ comes with two natural maps to X₀:

- The map πₖ is the map to the leftmost copy of X₀:
  \[π_k(x_1, \ldots, x_k) := π(x_1)\]
It is a covering map of degree $d^k$.

- The map $\phi_k$ is the map to the rightmost factor of $X_0$:
  \[ \phi_k(x_1, \ldots, x_k) := \phi(x_k). \]

It is a composition of lifts of $\phi$ to various covers of $X_0$.

We will also use the map $\phi$ to represent the entire virtual endomorphism.

This construction makes sense even if $\pi$ is not a covering map; in this generality, we are composing topological correspondences.

**7.2. Asymptotic stretch factors.** Now consider the case that $X$ is either a conformal surface $(\Sigma, \omega)$ or an elastic graph $(\Gamma, \alpha)$, with a virtual endomorphism $\pi, \phi : X_1 \Rightarrow X$ as above. If $X$ is a conformal surface, suppose also that $\phi$ is a topological embedding. Note that $X_k$ inherits the structure of a conformal surface or elastic graph from the covering map $\pi_k$.

**Definition 7.2.** The asymptotic stretch factor of the virtual endomorphism is

\[
\overline{SF}[\phi] = \lim_{n \to \infty} SF[\phi_n]^{1/n}.
\]

**Lemma 7.4.** The limit in Equation (7.3) exists.

**Proof sketch.** For graphs, sub-multiplicativity of SF (Proposition 6.4) and good behavior under covers (Proposition 6.37) show that

\[ SF[\phi_{k+1}] \leq SF[\phi_k] \cdot SF[\phi], \]

which implies that the sequence $SF[\phi_k]^{1/k}$ converges. (It is close to a decreasing sequence.) For surfaces, SF does not behave as well under covers, but we still have that if $\tilde{\phi}$ is a cover of $\phi$, then $SF[\phi] \leq SF[\tilde{\phi}] \leq \text{max}(1, SF[\phi])$ by Lemmas 6.11, 6.12, and 6.13; this is enough to show convergence. □

**Lemma 7.5.** The asymptotic stretch factor $\overline{SF}$ is independent of the conformal structure $\omega$ or elastic structure $\alpha$ used to define it. More generally, if $f : \Gamma \to \Gamma'$ is a homotopy equivalence of elastic graphs with homotopy inverse $g : \Gamma' \to \Gamma$, then

\[
\overline{SF}[f \circ \phi \circ g_1] = \overline{SF}[\phi],
\]

where $g_1 : \Gamma'_1 \to \Gamma_1$ is the lift of $g$ to the cover.

**Proof sketch.** Consider Equation (7.6) for graphs and let $\phi' = f \circ \phi \circ g_1$. We have $\phi'_n = f \circ \phi_n \circ g_n$. Therefore

\[ SF[\phi'_n] \leq (SF[f] SF[g_n]) SF[\phi_n], \]

with a similar inequality the other way. Since $SF[g_n] = SF[g]$ (Proposition 6.37), as $n \to \infty$ the contribution of the factor $SF[f] SF[g_n]$ to the limit goes to 1, so the asymptotic stretch factors are equal. The case of surfaces is similar. □

Thus we may speak about the asymptotic stretch factor of a virtual endomorphism $\pi, \phi : X_1 \Rightarrow X$ where $X$ is a topological graph or a surface, without reference to the conformal or elastic structure. When the covering map is trivial (i.e., the virtual endomorphism is an ordinary endomorphism), this recovers W. Thurston’s theory of pseudo-Anosov maps.
Proposition 7.7. If $\phi: \Sigma \to$ is a pseudo-Anosov self-homeomorphism of a surface $\Sigma$ (possibly with boundary), then $\overline{SF}[\phi]$ is the pseudo-Anosov constant of $\phi$, i.e., the exponential of the translation distance of the induced map on Teichmüller space.

Proof sketch. Follows from Theorem 3. \hfill \Box

Proposition 7.8. Let $\Gamma$ be a ribbon graph, and let $\pi, \phi: \Gamma_1 \to \Gamma$ be a virtual endomorphism of $\Gamma$ so that $\phi$ extends to a topological embedding $N\phi: N\Gamma_1 \hookrightarrow N\Gamma$. Then

$$\overline{SF}[\phi] = SF[N\phi].$$

Proof sketch. Proposition 6.38 relates stretch factors on graphs and on surfaces. The errors in the estimates disappear in the limit defining $\overline{SF}$, as in Lemma 7.5. \hfill \Box

Proposition 7.9. Let $\Sigma$ be a graph and $\pi, \phi: \Sigma_1 \to \Sigma$ be a virtual endomorphism of $\Sigma$. Then

$$SF[\phi] \leq SF[\phi].$$

Proof. Immediate from Proposition 6.4, Proposition 6.37, and the definition of $\overline{SF}$. \hfill \Box

Proposition 7.10. Let $\Sigma$ be a conformal surface and $\pi, \phi: \Sigma_1 \to \Sigma$ be a virtual endomorphism of $\Sigma$ with $\phi$ a conformal annular embedding. Then $\overline{SF}[\phi] < 1$.

Proof, version 1. By Theorem 5, $\tilde{SF}[\phi] < 1$. Then we have

$$SF[\phi_n] = SF[\phi_n \circ \tilde{\phi}] \leq SF[\phi_n] \tilde{SF}[\phi]$$

(where $\tilde{\phi}: X \to X$ is a cover of $\phi$), so

$$SF[\phi_n] \leq (\overline{SF}[\phi])^n$$

$$\overline{SF}[\phi] \leq SF[\phi].$$ \hfill \Box

Proof, version 2. Here we avoid Theorem 5.

Let $J(\phi)$ be the Julia set of $\phi$: the intersection of the images of $\phi_n(\Sigma_n)$. This set has measure 0. Suppose for simplicity that $\Sigma$ is connected; an Euler characteristic argument shows that $\Sigma$ is planar. Pick a conformal embedding of $\Sigma$ in $\mathbb{CP}^1$. For each $x \in J(\phi)$ and $n$ sufficiently large, $\phi_n(\Sigma_n)$ can be translated in $\Sigma \subset \mathbb{CP}^1$ to a map that misses $x$. By compactness of $J(\phi)$, for sufficiently large $n$ the homotopy class $[\phi_n]$ is conformally loose, so $\tilde{SF}[\phi_n] < 1$ by Proposition 6.17. This then implies $\overline{SF}[\phi] < 1$ as above. \hfill \Box

Proposition 7.11. For $\phi$ a virtual endomorphism of either an elastic graph or a conformal surface, if $\overline{SF}[\phi] < 1$, then for all sufficiently large $n$, $SF[\phi_n] < 1$.

Loosely speaking, Proposition 7.10 says that $\overline{SF}[\phi]$ detects conformal embeddings in covers. Since $\overline{SF}$ is also independent of the elastic or conformal structure used to define the stretch factor, this becomes a useful tool for studying rational maps.

Example 2.11 gives a case where the stretch factor necessarily drops under iteration, so $\overline{SF}[\phi] < SF[\phi]$. 
7.3. **Comparison to Lipshitz expansion.** There is another natural dynamical invariant of a (virtual) endomorphism of a graph: its (asymptotic) Lipshitz expansion, which we now compare with the asymptotic stretch factor.

**Definition 7.12.** For $\Gamma$ a metric graph and $\pi, \phi: \Gamma_1 \to \Gamma$ a virtual endomorphism of $\Gamma$, the Lipshitz energy of $\phi$ was defined in Equation (4.11). The **asymptotic expansion** of $\phi$ is

$$ \overline{\text{Lip}}[\phi] := \lim_{n \to \infty} \text{Lip}[\phi_n]^{1/n}. $$

This is independent of the metric graph used to define it (up to homotopy equivalence), by the argument in Lemma 7.5.

A **train track representative** for a free group automorphism $\psi: F_n \to F_n$ is an endomorphism $\phi$ of a graph $\Gamma$ inducing $\psi$ on $\pi_1$ so that $\text{Lip}[\phi^n] = \text{Lip}[\phi]^n$, i.e., so that $\overline{\text{Lip}}[\phi] = \text{Lip}[\phi]$. Bestvina and Handel’s theory of train tracks for free group automorphisms [BH92, Bes11] shows that most free group automorphisms have train track representatives. (More precisely, any irreducible automorphism has a train track representative.)

Given the similarity of the definitions, one might suspect these two quantities are related. Indeed, they are essentially the same in the case of endomorphisms that are not virtual.

**Proposition 7.13.** For $\phi: \Gamma \to \Gamma$ an endomorphism of a graph,

$$ \overline{SF}[\phi] = \overline{\text{Lip}}[\phi]^2. $$

**Proof sketch.** Suppose for simplicity that $\Gamma$ is trivalent and has $k$ edges each of length 1. Then, comparing Equations (5.11) and (5.13), we see that, for any curve $C$,

$$ \frac{1}{k} \ell[C]^2 \leq \text{EL}[C] \leq \ell[C]^2. $$

We deduce, from Equation (4.13) and the definition of stretch factor, that

$$ (7.14) \quad \frac{1}{k} \text{Lip}[\phi]^2 \leq \overline{SF}[\phi] \leq k \text{Lip}[\phi]^2. $$

The error factors in these inequalities go away in the limits defining $\overline{SF}$ and $\overline{\text{Lip}}$. □

**Remark 7.15.** Proposition 7.13 may be compared to the fact that the absolute value of the largest eigenvalue of a linear operator on a finite-dimensional space equals the asymptotic growth rate of the norm of any vector, independent of the choice of norm. In this dictionary, EL is like the (square of the) $L^2$ norm while $\ell$ is like the $L^1$ norm.

For a **virtual** endomorphism $\pi, \phi: \Gamma_1 \to \Gamma$, the quantities $\overline{SF}$ and $\overline{\text{Lip}}$ are in general different. By Equation (6.34), we have the rather weak inequality

$$ (7.16) \quad \overline{\text{Lip}}[\phi] \leq \overline{SF}[\phi]. $$

See Section 8.6 for a comparison of how these two quantities relate to dynamics on $S^2$.

**Remark 7.17.** The reason it is impossible to find train-track representatives in cases like Example 2.11 is that the same vertex in $\Gamma$ appears multiple times in $\Gamma_1$, and it is impossible to arrange for all of these vertices to simultaneously obey the train-track requirements.
7.4. Extensions and questions. There are many questions raised by this theory. First of all, in order to cover the general case of rational maps we need to extend the definitions to orbifold fundamental groups.

**Problem 7.18.** Extend the definition of $\text{SF}$ to automorphisms of pointed graphs, or to virtual endomorphisms of orbifolds.

**Problem 7.19.** More generally, define and study $\text{SF}$ for a virtual endomorphism of an arbitrary group.

We can also ask about properties of the result.

**Question 7.20.** Is $\text{SF}[\phi]$ algebraic?

**Problem 7.21.** Give an algorithm for computing $\text{SF}[\phi]$, either approximately or exactly.

8. Rational maps

Finally, we come to the original goal of this work, the study of branched self-covers of the sphere and when they are equivalent to rational maps.

**Definition 8.1.** Fix a finite set $P$ of points in a sphere $S^2$. A branched self-cover $f: (S^2, P) \to S^2$ is a map so that $f(P) \subset P$ and so that $f$ is a covering map when restricted to $S^2 \setminus f^{-1}(P) \to S^2 \setminus P$. Two branched self-covers are equivalent if they are related by homotopy relative to $P$ and by conjugacy. (Without conjugacy, we could not change the set $P$.) A branched self-cover is rational if it is equivalent to a rational map on $\mathbb{C}P^1$, which is necessarily post-critically finite. ($P$ contains the post-critical set, but may be larger.)

We will assume there is a branch point in each cycle of $P$. For rational maps, implies $f$ is hyperbolic.

8.1. Characterizing rational maps conformally. We first give a characterization of rational maps in terms of surfaces.

**Definition 8.2.** A surface spine for $S^2 \setminus P$ is a surface $\Sigma$ (necessarily planar) together with a topological embedding $i: \Sigma \hookrightarrow S^2 \setminus P$ so that complement of the image consists of one punctured disk for each point in $P$, i.e., so that $S^2 \setminus P$ deformation retracts onto $i(\Sigma)$.

**Definition 8.3.** Let $f: (S^2, P) \to S^2$ be a branched self-cover and $\Sigma$ be a surface equipped with an embedding $i: \Sigma \hookrightarrow S^2 \setminus P$. Then the inverse image $f^{-1}(\Sigma)$ is the pull-back in the diagram

$$
\begin{array}{ccc}
  f^{-1}(\Sigma) & \xrightarrow{\pi} & \Sigma \\
  \downarrow & & \downarrow \ i \\
  S^2 \setminus f^{-1}(P) & \xrightarrow{f} & S^2 \setminus P.
\end{array}
$$

There are two natural maps:

- An embedding $f^{-1}(\Sigma) \hookrightarrow S^2 \setminus f^{-1}(P) \subset S^2 \setminus P$. 

A covering map \( \pi: f^{-1}(\Sigma) \to \Sigma \). Since it is a covering map, \( f^{-1}(\Sigma) \) inherits all the structure of \( \Sigma \).

If \( \Sigma \) is a surface spine for \( S^2 \setminus P \), then the embedding \( f^{-1}(\Sigma) \to S^2 \setminus P \) is isotopic to a topological embedding \( \phi: f^{-1}(\Sigma) \to \Sigma \). We therefore have a virtual endomorphism of \( \Sigma \) and can apply the technology of Section 7.

**Theorem 8.4.** Let \( f: (S^2, P) \to (S^2, P) \) be a branched self-cover with a branch point in each cycle. Then \( f \) is equivalent to a rational map iff there is a conformal surface spine \( \Sigma \) for \( S^2 \setminus P \) so that in the corresponding virtual endomorphism, \( \phi \) is homotopic to a strict conformal embedding \( \phi: f^{-1}(\Sigma) \to \Sigma \).

The above characterization of rational maps appears to have been folklore in the community for some time.

**Proof sketch.** If \( f: (\mathbb{C}P^1, P) \to (\mathbb{C}P^1, P) \) is a rational map, then we may take \( \Sigma \) to be a neighborhood of the Julia set \( J(f) \) (filling in all disks that do not contain critical points). Since the dynamics is super-attracting to \( P \) on the complement of \( J(f) \) than \( \Sigma \), and by taking \( \Sigma \) suitably nice (e.g., using level sets of the potential function) we can arrange for \( f^{-1}(\Sigma) \subset \Sigma \) to be a strict conformal embedding.

The converse direction is a special case of [CPT, Theorem 5.2]. The technique, quasi-conformal surgery, goes back to Douady and Hubbard [DH85]. □

**Remark 8.5.** In fact, Theorem 8.4 is true if the assumption that the conformal embedding is strict is dropped. With that modification, it also extends to more general branched self-covers. This strengthening is not needed for the main theorem.

**8.2. Characterizing rational maps using graphs.** In the setting of Theorem 7, we can consider the asymptotic stretch factor of the virtual endomorphism. Theorem 8.4 and Propositions 7.10 and 7.11 then tell us the following.

**Proposition 8.6.** Let \( f: (S^2, P) \to (S^2, P) \) be a branched self-cover with a branch point in each cycle in \( P \) and let \( \pi, \phi: \Sigma_1 \to \Sigma \) be the corresponding virtual endomorphism. Then \( f \) is equivalent to a rational map iff \( \text{SF}[\phi] < 1 \).

By the equality of the asymptotic stretch factors for graphs and for surfaces (Proposition 7.8), we immediately get the following more precise version of Theorem 1.

**Theorem 8.** Let \( [f]: (S^2, P) \to (S^2, P) \) be a branched self-cover of the sphere relative to a finite number of points \( P \subset S^2 \), with a branch point in each cycle in \( P \). For \( \Gamma \) any spine of \( S^2 \setminus P \), let \( \pi_\Gamma, \phi_\Gamma: \Gamma_1 \to \Gamma \) be the corresponding virtual endomorphism. Then the following conditions are equivalent.

1. The branched self-cover \( f \) is equivalent to a rational map.
2. For any spine \( \Gamma \) for \( S^2 \setminus P \), we have \( \text{SF}[\phi_\Gamma] < 1 \).
3. There is some elastic graph spine \( \Gamma \) for \( S^2 \setminus P \) and some integer \( n > 0 \) so that \( \text{Emb}[\phi_{\Gamma,n}] < 1 \).
4. For every elastic graph spine \( \Gamma \) for \( S^2 \setminus P \) and every sufficiently large \( n \), we have \( \text{Emb}[\phi_{\Gamma,n}] < 1 \).

**Proof sketch.** The proof has essentially already been given earlier in the paper. The diagram below summarizes the chain of implications between (1), (2), and (3). Here,
Σ is a (conformal) surface spine for $S^2 \setminus P$. There are various side conditions for which you should see the referenced theorems and propositions.

\[
\begin{align*}
\text{f rational} & \quad \text{Thm. 8.4} \\
\exists \Sigma: f^{-1}(\Sigma) \text{ strictly conformally embeds in } \Sigma & \quad \exists \Gamma, n: \text{Emb}[\phi_{\Gamma,n}] < 1 \\
\text{Thm. 4} & \quad \text{Prop. 7.10, See below} \\
\exists \Sigma: \text{SF}[\phi_{\Sigma}] < 1 & \quad \exists \Gamma, n: \text{SF}[\phi_{\Gamma,n}] < 1 \\
\text{Prop. 7.11} & \quad \text{Prop. 7.11} \\
\text{SF}[\phi_{\Sigma}] < 1 < \text{Prop. 7.8} & \quad \text{SF}[\phi_{\Gamma}] < 1.
\end{align*}
\]

Lemma 7.5 says that the asymptotic stretch factor $\text{SF}$ is independent of the conformal structure on $\Sigma$ or elastic structure on $\Gamma$, giving the equivalence to statement (4).

The implication

\[ \text{SF}[\phi_{\Sigma}] < 1 \implies \exists \Sigma: \text{SF}[\phi_{\Sigma}] < 1 \]

is not direct. Proposition 7.11 gives the implication

\[ \text{SF}[\phi_{\Sigma}] < 1 \implies \exists \Sigma, n: \text{SF}[\phi_{\Sigma,n}] < 1, \]

which by Theorem 8.4 implies that $f^{on}$ is equivalent to a rational map. To then conclude that $f$ is a rational map and that we don’t need to iterate to find a surface spine $\Sigma$ with SF[$\phi_{\Sigma}$] < 1, we need to know a little bit more about the geometry. For instance, it suffices to know that, except in Lattès examples, some power of the backwards iteration map on Teichmüller space is contracting. This is the only place that this argument relies on the original characterization of rational maps [DH93].

The virtual endomorphism $\phi: f^{-1}(\Gamma) \to \Gamma$ is going in essentially the reverse direction to $f$: While $f$ is expanding on the Julia set, $\phi$ is contracting (in an appropriate sense). Therefore, for $f$ a branched self-cover with a branch point in each cycle and $\phi$ the associated virtual endomorphism of a spine, we define

\[ \text{SF}[f] := \text{SF}[\phi]^{-1}. \]

This quantity is greater than one when $f$ is rational.

8.3. Polynomials. One important special case is that of polynomials, where Theorem 8 recovers known results, which we briefly review.

Definition 8.8. A topological polynomial is a branched self-cover $[f]: (S^2, P) \to$ of degree $d$ with one fixed point $\infty \in P$ which is branched of degree $d$. If such a map is equivalent to a rational map, the rational map is a polynomial.
Definition 8.9. For a post-critically finite polynomial $p$ with post-critical set $P \cup \{\infty\} \subset \mathbb{CP}^1$, the filled Julia set $K(p)$ is the set of points bounded under iteration. It is the union of the Julia set $J(p)$ and the attracting basin around each finite Fatou point, points that eventually maps to a periodic cycle with a branch point. The Hubbard tree $T_P \subset K(p)$ is essentially the spanning tree of $P$ within $K(p)$. When $T_P$ intersects a Fatou component, it is required to be a union of rays in the canonical coordinates. See [DH85].

Now define the Hubbard graph $G_P \subset J(p)$ to be the union of
- the boundary of the Fatou component containing each Fatou point in $P$,
- the Julia points of $P$, and
- the minimal tree connecting the above points in $K(p)$, taking unions of rays in coordinates of Fatou components as before.

$T_P$ is forward invariant under the polynomial $p$. As such, we can consider dynamics on it. To make it a complete invariant of the map, we can decorate $T_P$ with some additional angle data at Fatou points [DH84, Section 6.1.2]; this is close to what is needed to reconstruct the graph $G_P$ from the map on $T_P$. A abstract Hubbard tree is a tree with an endomorphism, marked with this additional data.

Definition 8.10. An endomorphism $f$ of a Hubbard tree is Julia expanding if it admits a metric for which $f$ does not decrease distances and strictly increases the distance between any two Julia points.

The Hubbard graph $G_P$, on the other hand, is not in general forward invariant, since near a point in $f^{-1}(P) \setminus P$ the graph $G_P$ may go through a Fatou component but its image does not. It does have a natural virtual endomorphism, which can be reconstructed from the abstract Hubbard tree.

Lemma 8.11. An endomorphism of an abstract Hubbard tree is Julia expanding iff the corresponding virtual endomorphism of the abstract Hubbard graph has embedding energy less than 1.

Theorem 8.12 (Poirier [Poi09,Poi10]). An abstract Hubbard tree $T_P$ is realizable as the Hubbard tree of a polynomial iff it is Julia expanding.

In the presence of expanding dynamics, we can consider the entropy.

Definition 8.13. The core entropy $h(f)$ of a post-critically finite polynomial $f$ is the topological entropy of the map on the Hubbard tree $T_P$.

Proposition 8.14. Let $f$ be a post-critically finite quadratic polynomial. If $f$ is dendritic (i.e., each critical point is strictly pre-periodic), then
\[ \overline{SF}[f] = h(f). \]

Here we are using an extension of the theory of stretch factors, etc., to the non-hyperbolic case, as mentioned in Problem 8.23 below.

If $f$ is hyperbolic (i.e., each critical point end up in a cycle with at least one critical point) and the critical cycles have degree $d_i$ and length $n_i$ (for $i = 1, \ldots, k$) then
\[ \overline{SF}[f] \geq \max_{1 \leq i \leq k} \frac{\log d_i}{n_i}. \]
This last inequality is often an inequality and often gives answers larger than $h(f)$.

The fact that $f$ is realized as a polynomial iff the entropy on a topological Hubbard graph is positive is essentially a corollary of Poirier’s Theorem 8.12.

8.4. **Comparison to annular obstruction.** As mentioned in the introduction, W. Thurston in 1982 gave a different characterization of rational maps among topological branched self-covers, as explained by Douady and Hubbard. Instead of finding an object that guarantees that the map is rational, he finds an obstruction that guarantees it is not rational. We can write down that obstruction in our language as follows.

**Definition 8.15.** An *annular* elastic graph is one that consists only of circle components. Let $A$ be an annular elastic graph embedded in an orbifold $\Sigma$. Then the *join* of $A$ is the elastic graph $\text{Join}(A)$ obtained by

- deleting all components of $A$ that bound a disk with at most one orbifold point, and then
- merging all parallel components of $A$ by harmonically adding their elastic constants.

*Harmonically adding* the elastic constants means taking two parallel components $a_1$ and $a_2$ and replacing them with a new component $a_3$ with

$$\frac{1}{\alpha(a_3)} = \frac{1}{\alpha(a_1)} + \frac{1}{\alpha(a_2)}.$$

This corresponds to stitching together two side-by-side conformal annuli in the most efficient way possible.

**Theorem 8.16** (W. Thurston, Douady-Hubbard [DH93]). Let $f : (S^2, P) \Leftarrow$ be a topological branched self-cover that is not a Lattés example. Let $S^2_f$ be the orbifold of $f$. Then $f$ is equivalent to a rational map iff there is no annular elastic graph $A$ in $S^2_f$ and map $\phi : A \to \text{Join}(f^{-1}(A))$ compatible with the maps to $S^2 \setminus P$ so that $\text{Emb}(\phi) \leq 1$.

**Remark** 8.17. The usual formulation of Theorem 8.16 refers to the maximum eigenvalue of a matrix constructed out of $A$ considered as a multi-curve (with no extra structure). The above formulation is easily seen to be equivalent.

Where our Theorem 8 looks for a strict conformal embedding of graphs

$$f^{-1}(\Gamma) \hookrightarrow \Gamma,$$

the older Theorem 8.16 looks for a (not necessarily strict) conformal embedding the other direction

$$A \hookrightarrow \text{Join}(f^{-1}(A)).$$

There is an easy argument that both conditions cannot simultaneously hold (as implied by the theorems). If both embeddings existed, scale $A$ so that there is a tight conformal embedding $A \hookrightarrow \Gamma$ (i.e., one with stretch factor equal to 1), and consider
the square of inclusions

\[ \text{Join}(f^{-1}(A)) \quad \xrightarrow{\quad \downarrow \quad} \quad A \]

(8.18)

\[ f^{-1}(\Gamma) \quad \xrightarrow{\quad \downarrow \quad} \quad \Gamma \]

The map \( f^{-1}(\Gamma) \hookrightarrow \Gamma \) is a strict conformal embedding but the map \( A \hookrightarrow \Gamma \) is not, a contradiction.

It is also worth noting that Theorem 8 is technically easier than the existing proof of Theorem 8.16. The key analytic point is contained in Theorem 8.4, which is relatively soft.

8.5. Comparison to domination of weighted arc diagrams. Suppose we are given a virtual endomorphism of surfaces \( \pi, \phi : \Sigma_1 \Rightarrow \Sigma_0 \), with \( \pi \) a covering map and \( \phi \) a topological embedding. Then the associated dynamical Teichmüller space is

\[ \text{Teich}(\pi, \phi) := \{ S \in \text{Teich}(\Sigma_0) \mid \pi^* S \text{ conformally embeds in } S \text{ in the homotopy class } [\phi] \} \].

Here, \( \text{Teich}(\Sigma_0) \) is the finite-dimensional space of conformal structures on the interior of \( \Sigma_0 \). (The conformal structures are allowed to have removable singularities or not; that is, the hyperbolic structures can have parabolic or hyperbolic monodromy around the boundary components.) Then Theorem 8.4 and the strengthening in Remark 8.5 say that the virtual endomorphism comes from a rational map iff \( \text{Teich}(\pi, \phi) \) is non-empty.

Jeremy Kahn has studied the topology of \( \text{Teich}(\pi, \phi) \) in the context of renormalization of polynomials, and has results in terms of weighted arc diagrams [Kah06]. A weighted arc diagram is a weighted collection \( X \) of arcs that connect boundary components of a surface \( \Sigma \) with boundary. If \( X \) is filling (the complementary components are disks), it can be thought of as dual to an elastic graph \( \Gamma(X) \) embedded in \( \Sigma \), with one vertex per component of \( \Sigma \setminus X \) and one edge crossing each arc in \( X \). The elastic length of an edge of \( \Gamma(X) \) is equal to the weight on the arc it crosses.

Now suppose that you have two surfaces, \( \Sigma_1 \subset \Sigma_0 \). There if \( X \) is a weighted arc diagram on \( \Sigma_1 \) and \( Y \) is a weighted arc diagram on \( \Sigma_0 \), there is a notion of \( X \) dominating \( Y \), written \( X \twoheadrightarrow Y \). We do not repeat the definition here (see [Kah06, Section 3.6]), but, in the filling case, \( X \) dominates \( Y \) iff there is a map \( \phi : \Gamma(X) \to \Gamma(Y) \) with \( \text{Emb}(\phi) \leq 1 \) that is compatible up to homotopy with the inclusion \( \Sigma_1 \subset \Sigma_0 \).

Kahn used this notion of domination to control certain infinitely renormalizable polynomials (studying how conformal structures can degenerate). In terms of Teichmüller spaces, we have the following:

Proposition 8.19 (Kahn, personal communication). If \( \text{Teich}(\pi, \phi) \) is not compact, then there is a weighted arc diagram \( X \) on \( \Sigma_0 \) so that \( \pi^* X \twoheadrightarrow X \).

Conjecture 8.20 (Kahn, personal communication). If there is a weighted arc diagram \( X \) on \( \Sigma_0 \) so that \( \pi^* X \twoheadrightarrow X \), then \( \text{Teich}(\pi, \phi) \) is not compact.

We can prove a weaker version of Conjecture 8.20:
Corollary 8.21. If there is a weighted arc diagram $X$ on $\Sigma_0$ so that $\pi^*X$ strictly dominates $X$ and $X$ fills $\Sigma_0$, then $\Teich(\pi, \phi)$ is not compact.

Proof. The conditions are equivalent to saying that $\Gamma(\pi)$ is a spine for $\Sigma_0$ and $\Emb[\pi^*\Gamma(X) \to \Gamma(X)] < 1$. Then Theorem 7 says that $N_t\Gamma(X) \in \Teich(\pi, \phi)$ for $t$ sufficiently small; this gives an explicit sequence exhibiting the non-compactness of $\Teich(\pi, \phi)$. □

Corollary 8.22. If the virtual endomorphism $\pi, \phi : \Sigma_1 \to \Sigma$ comes from a topological branched self-cover $f : (S^2, P) \to (S^2, P)$ with at least one branch point in each cycle, then $\Teich(\pi_n, \phi_n)$ is not compact for $n$ sufficiently large.

Proof. Theorem 8 says that for $n$ large enough, there is an elastic graph $\Gamma$ so that $\Emb[r, f_n^*(\Gamma) \to \Gamma] < 1$, which by Corollary 8.21 implies $\Teich(\pi_n, \phi_n)$ is not compact. □

In fact, the techniques can be strengthened to show that as long as $\Teich(\pi, \phi)$ is more than a single point, then $\Teich(\pi_n, \phi_n)$ is eventually non-compact.

In addition, Kahn has communicated examples where $\Teich(\pi, \phi)$ is compact (but more than one point), thus showing that iteration is necessary in Theorem 8.

8.6. Comparison to Lipshitz expansion. In Section 7.3 we remarked that there are two analytic quantities associated a virtual endomorphism of a graph: the asymptotic stretch factor $SF$ and the asymptotic Lipshitz constant $\overline{\Lip}$. In the setting of Theorem 8, what are the graphs for which $\overline{\Lip}[f^{-1}(\Gamma) \to \Gamma] < 1$? Since $\overline{\Lip}[\phi] \leq SF[\phi]$ (Equation (7.16)), this is a weaker condition. In fact, $\overline{\Lip}[\phi] < 1$ can be taken as a definition of when associated virtual endomorphism of $\pi_1(S^2 \setminus P)$ is contracting in the sense of Nekrashevych [Nek05, Nek11, Nek14].

Bonk and Meyer [BM10] and Haïssinsky and Pilgrim [HP12] studied essentially this condition, in the opposite situation to the present paper: they considered the case when there are no branch points in any cycle in $P$. (For rational maps, this means that the Julia set is the whole sphere.) In our language, they showed more-or-less that for a branched self-cover $f$ with no branch points in periodic cycles with associated virtual endomorphism $\phi$, the stretch factor $\overline{\Lip}[\phi] < 1$ if $f$ has a topological representative as a map on $S^2$ that is uniformly expanding. Any rational map has such a uniformly expanding representative, and many other maps do as well.

8.7. Extensions and questions. The obvious question is to extend Theorem 8 away from the restricted setting.

Problem 8.23. Extend Theorem 8 to general topological branched self-covers, dropping the condition that there be a critical point in each cycle.

It appears that the theory will extend without much problem to the much larger family of maps which have at least one critical point in one cycle in $P$. (For rational maps, these are maps where the Julia set is not the entire sphere.) This requires extending the theory for graphs and surfaces to allow dealing with marked points, as briefly mentioned in Section 3.4.

For the remaining case, when there are no branch points in any cycles in $P$ and the Julia set is the whole sphere if the map is rational, there is no clear analogue of the
conformal characterization in Theorem 8.4 and so it is not clear what the statement would be.

**Question 8.24.** In Theorem 8, does \( n = 1 \) suffice? That is, for \( f \) as given there, is there always an elastic graph \( \Gamma \) so that there is a strict conformal embedding of \( f^{-1}(\Gamma) \) into \( \Gamma \)?

The answer to Question 8.24 is “yes” for polynomials and for many examples, and Theorem 8.4 is a corresponding statement for conformal surfaces. Nevertheless, the answer to Question 8.24 is “no” in general. See Section 8.5.

**Problem 8.25.** In Theorem 8, how does the minimal necessary \( n \) grow, as a function of the degree and the size of \( P \)?

**Problem 8.26.** Study the algorithmic complexity of finding a strict conformal embedding \( \phi: f^{-n}(\Gamma) \to \Gamma \) when \( f \) is equivalent to a rational map. In particular, does this give a polynomial-sized certificate that \( f \) is rational?

Presumably a solution to Problem 8.26 would first require a solution to Problem 8.25.

**Problem 8.27.** Study \( SF[f] \) as \( f \) varies over rational maps. How does it compare to known invariants?

Question 7.20 is one starting question. The connection between \( SF \) and the entropy on the Hubbard graph described in Section 8.3 is relevant, but in general there is no reasonable forward dynamics on \( \Gamma \).

Finally, Theorems 8 and 8.16 give opposite combinatorial conditions. These should presumably be related.

**Problem 8.28.** Prove combinatorially that, for any virtual endomorphism \( \phi: \tilde{\Gamma} \to \Gamma \) of a graph, either \( SF[\phi] < 1 \), or there is an annular obstruction in the sense of Theorem 8.16.

A solution to Problem 8.28 would give an alternate proof of the main Theorem 8, using W. Thurston’s Theorem 8.16.

Another family of problems comes from applying Theorem 8: Prove that certain maps are rational by finding explicit maps \( \phi: f^{-1}(\Gamma) \to \Gamma \) with \( \text{Emb}(\phi) < 1 \). We give one sample problem.

**Problem 8.29.** Rees and Tan Lei [Tan92], and later Shishikura [Shi00] showed that two quadratic polynomials are topologically mateable iff they do not lie in conjugate limbs of the Mandelbrot set. Reprove this with the techniques of this paper (See Examples 2.6 and 2.8 for one approach.) More generally, find a criterion for cubic or higher-degree polynomials to be mateable; this is likely to be more difficult, given the cubic example of Shishikura and Tan Lei without Levy cycles [ST00].

**Warning** 8.30. A topological mating is not the naive (formal) mating obtained by gluing two polynomials at infinity. See the survey by Buff et al. [BEK+12].
References

[Ago11] Ian Agol, *Ideal triangulations of pseudo-Anosov mapping tori*, Topology and geometry in dimension three, Contemp. Math., vol. 560, Amer. Math. Soc., Providence, RI, 2011, pp. 1–17.

[BBY12] Sylvain Bonnot, Mark Braverman, and Michael Yampolsky, *Thurston equivalence to a rational map is decidable*, Mosc. Math. J. 12 (2012), no. 4, 747–763, arXiv:1009.5713.

[BEK+12] Xavier Buff, Adam L. Epstein, Sarah Koch, Daniel Meyer, Kevin Pilgrim, Mary Rees, and Tan Lei, *Questions about polynomial matings*, Ann. Fac. Sci. Toulouse Math. (6) 21 (2012), no. 5, 1149–1176.

[Bes11] Mladen Bestvina, *A Bers-like proof of the existence of train tracks for free group automorphisms*, Fund. Math. 214 (2011), no. 1, 1–12, arXiv:1001.0325.

[BH92] Mladen Bestvina and Michael Handel, *Train tracks and automorphisms of free groups*, Ann. of Math. (2) 135 (1992), no. 1, 1–51.

[BM10] Mario Bonk and Daniel Meyer, *Expanding Thurston maps*, Preprint, 2010, arXiv:1009.3647.

[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*, The mathematical legacy of Wilhelm Magnus: groups, geometry and special functions (Brooklyn, NY, 1992), Contemp. Math., vol. 169, Amer. Math. Soc., Providence, RI, 1994, pp. 133–212.

[CP] Guizhen Cui, Wenjuan Peng, and Tan Lei, *Renormalization and wandering Jordan curves of rational maps*, Preprint, arXiv:1403.5024, 2014.

[DH84] Adrien Douady and John H. Hubbard, *Étude dynamique des polynômes complexes. Partie I*, Publications Mathématiques d’Orsay, vol. 84–2, Université de Paris-Sud, Département de Mathématiques, Orsay, 1984, available in English as *Exploring the Mandelbrot set*. The Orsay notes from http://math.cornell.edu/~hubbard.

[DH85] Adrien Douady and John H. Hubbard, *On the dynamics of polynomial-like mappings*, Ann. Sci. École Norm. Sup. (4) 18 (1985), no. 2, 287–343.

[DH93] *A proof of Thurston’s topological characterization of rational functions*, Acta Math. 171 (1993), no. 2, 263–297.

[Duf62] R. J. Duffin, *The extremal length of a network*, J. Math. Anal. Appl. 5 (1962), 200–215.

[EF01] J. Eells and B. Fuglede, *Harmonic maps between Riemannian polyhedra*, Cambridge Tracts in Mathematics, vol. 142, Cambridge University Press, Cambridge, 2001, With a preface by M. Gromov.

[FM11] Stefano Francaviglia and Armando Martino, *Metric properties of outer space*, Publ. Mat. 55 (2011), no. 2, 433–473, arXiv:0803.0640.

[Gar84] Frederick P. Gardiner, *Measured foliations and the minimal norm property for quadratic differentials*, Acta Math. 152 (1984), no. 1-2, 57–76.

[Gar87] *Teichmüller theory and quadratic differentials*, Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1987, A Wiley-Interscience Publication.

[GM91] Frederick P. Gardiner and Howard Masur, *Extremal length geometry of Teichmüller space*, Complex Variables Theory Appl. 16 (1991), no. 2-3, 209–237.

[Ham09] Ursula Hamenstädt, *Geometry of the mapping class groups. I. Boundary amenability*, Invent. Math. 175 (2009), no. 3, 545–609.

[HM79] John Hubbard and Howard Masur, *Quadratic differentials and foliations*, Acta Math. 142 (1979), no. 3-4, 221–274.

[HP12] Peter Haïssinsky and Kevin M. Pilgrim, *An algebraic characterization of expanding Thurston maps*, J. Mod. Dyn. 6 (2012), no. 4, 451–476, arXiv:1204.3214.

[Iof75] M. S. Ioffe, *Extremal quasiconformal imbeddings of Riemann surfaces*, Sibirsk. Mat. Ž. 16 (1975), no. 3, 520–537, 644.

[Kah06] Jeremy Kahn, *A priori bounds for some infinitely renormalizable quadratics: I. Bounded primitive combinatorics*, Preprint ims06-05, 2006, arXiv:math/0609045v2.
[Ker80] Steven P. Kerckhoff, *The asymptotic geometry of Teichmüller space*, Topology 19 (1980), no. 1, 23–41.

[LM97] Mikhail Lyubich and Yair Minsky, *Laminations in holomorphic dynamics*, J. Differential Geom. 47 (1997), no. 1, 1–195.

[MS84] Albert Marden and Kurt Strebel, *The heights theorem for quadratic differentials on Riemann surfaces*, Acta Math. 153 (1984), no. 3-4, 153–211.

[Nek05] Volodymyr Nekrashevych, *Self-similar groups*, Mathematical Surveys and Monographs, vol. 117, American Mathematical Society, Providence, RI, 2005.

[Nek11] , *Iterated monodromy groups*, Groups St Andrews 2009 in Bath. Volume 1, London Math. Soc. Lecture Note Ser., vol. 387, Cambridge Univ. Press, Cambridge, 2011, pp. 41–93, arXiv:math/0312306.

[Nek14] , *Combinatorial models of expanding dynamical systems*, Ergod. Thy. & Dynam. Sys. 34 (2014), no. 3, 938–985, arXiv:0810.4936.

[Pil01] Kevin M. Pilgrim, *Canonical Thurston obstructions*, Adv. Math. 158 (2001), no. 2, 154–168.

[Poi09] Alfredo Poirier, *Critical portraits for postcritically finite polynomials*, Fund. Math. 203 (2009), no. 2, 107–163.

[Poi10] , *Hubbard trees*, Fund. Math. 208 (2010), no. 3, 193–248.

[PP87] Athanase Papadopoulos and Robert C. Penner, *A characterization of pseudo-Anosov foliations*, Pacific J. Math. 130 (1987), no. 2, 359–377.

[PT98] Kevin Pilgrim and Tan Lei, *Combining rational maps and controlling obstructions*, Ergodic Theory Dynam. Systems 18 (1998), no. 1, 221–245.

[Roc70] R. Tyrell Rockafellar, *Convex analysis*, Princeton University Press, 1970.

[Sch93] Oded Schramm, *Square tilings with prescribed combinatorics*, Israel J. Math. 84 (1993), no. 1-2, 97–118.

[Shi00] Mitsuhiro Shishikura, *On a theorem of M. Rees for matings of polynomials*, The Mandelbrot set, theme and variations, London Math. Soc. Lecture Note Ser., vol. 274, Cambridge Univ. Press, Cambridge, 2000, pp. 289–305.

[ST00] Mitsuhiro Shishikura and Tan Lei, *A family of cubic rational maps and matings of cubic polynomials*, Experiment. Math. 9 (2000), no. 1, 29–53.

[Tan92] Tan Lei, *Matings of quadratic polynomials*, Ergodic Theory Dynam. Systems 12 (1992), no. 3, 589–620.

[Thu] Dylan Thurston, *On geometric intersection of curves in surfaces*, draft available at [http://math.columbia.edu/~dpt/DehnCoordinates.pdf](http://math.columbia.edu/~dpt/DehnCoordinates.pdf).

[Wol98] Michael Wolf, *Measured foliations and harmonic maps of surfaces*, J. Differential Geom. 49 (1998), no. 3, 437–467.