1.3 Rubber bands and connectivity

Rubber band representations can be related to graph connectivity, and can be used to give a test for k-connectivity of a graph (Linial, Lovász and Wigderson [2]).

1.3.1 Degeneracy: essential and non-essential

We start with a discussion of what causes degeneracy in rubber band embeddings. Consider the two graphs in Figure 1.1. It is clear that if we nail the nodes on the convex hull, and then let the rest find its equilibrium, then there will be a degeneracy: the grey nodes will all move to the same position. However, the reasons for this degeneracy are different: In the first case, it is due to symmetry; in the second, it is due to the node that separates the grey nodes from the rest, and thereby pulls them onto itself.

One can distinguish the two kinds of degeneracy as follows: In the first graph, the strengths of the rubber bands must be strictly equal; varying these strengths it is easy to break the symmetry and thereby get rid of the degeneracy. However, in the second graph, no matter how we change the strengths of the rubber bands (as long as they remain positive), the grey nodes will always be pulled together into one point.

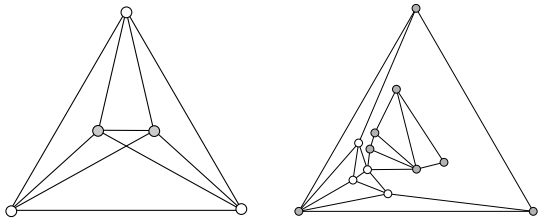

Figure 1.1: Two reasons why two nodes end up on top of each other: symmetry, or a separating node

Figure 1.2 illustrates a bit more delicate degeneracy. In all three pictures, the grey points end up collinear in the rubber band embedding. In the first graph, the reason is symmetry again. In the second, there is a lot of symmetry, but it does not explain why the three grey nodes are collinear in the equilibrium. (It is not hard to argue though that they are collinear: a good exercise!) In the third graph (which is not drawn in its equilibrium position, but before it) there are two nodes separating the grey nodes from the nailed nodes, and the grey nodes will end up on the segment connecting these two nodes. In the first two cases, changing the strength of the rubber bands will pull the grey nodes off the line; in the third, this does not happen.

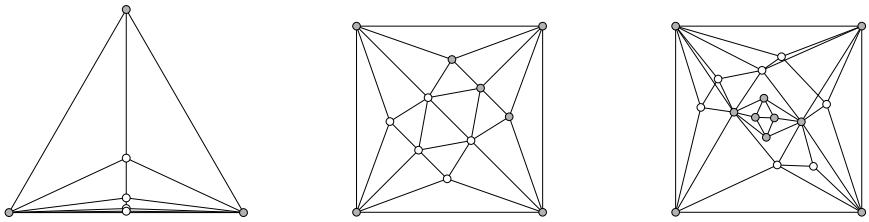

Figure 1.2: Three reasons why three nodes can end up collinear: symmetry, just accident, or a separating pair of nodes

1.3.2 Connectivity and degeneracy

Our goal is to prove that essential degeneracy in a rubber band embedding is always due to low connectivity. We start with the easy direction, formalizing the examples from the previous section.

Lemma 1.3.1 *Let $G = (V, E)$ be a graph and $S, T \subseteq V$. Then for every rubber band representation $\mathbf{x}$ of G with S nailed,* $\mathrm{rk}(\mathbf{x}(T)) \leq \kappa(S, T)$.

Proof. There is a subset $U \subseteq V$ with $|U| = \kappa(S, T)$ such that $V \setminus U$ contains no (S, T)-paths. Let W be the union of connected components of $G \setminus U$ containing a vertex from T. Then $\mathbf{x}$, restricted to W, gives a rubber band representation of $G[W]$ with boundary U. Clearly $\mathbf{x}(W) \subseteq \mathrm{conv}(\mathbf{x}(U))$, and so

$$\mathrm{rk}(\mathbf{x}(T)) \leq \mathrm{rk}(\mathbf{x}(W)) = \mathrm{rk}(\mathbf{x}(U)) \leq |U| = \kappa(S, T).$$

□

The Lemma gives a lower bound on the connectivity between two sets S and T. The following theorem asserts that if we take the best convex representation, this lower bound is tight:

Theorem 1.3.2 *Let $G = (V, E)$ be a graph and $S, T \subseteq V$ with $\kappa(S, T) = d+1$. Then G has a rubber band representation in $\mathbb{R}^d$ with S nailed such that* $\mathrm{rk}(\mathbf{x}_T) = d+1$.

Corollary 1.3.3 *Let $G = (V, E)$ be a graph, $d \geq 1$ and $S \subseteq V$. Then G has a rubber band representation in $\mathbb{R}^d$ in general position with S nailed if and only if no node of G can be separated from S by fewer than $d+1$ nodes.*

Corollary 1.3.4 *A graph G is k-connected if and only if for every $S \subseteq V$ with $|S| = k$, G has a rubber band representation in $\mathbb{R}^{k-1}$ in general position with S nailed.*

To prove Theorem 1.3.2, we choose generic edge weights.

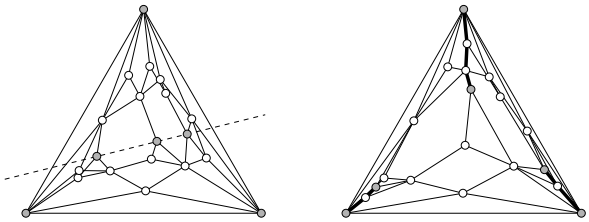

Figure 1.3: Three nodes accidentally on a line. Strengthening the edges along the three paths pulls them apart.

Theorem 1.3.5 *Let $G = (V, E)$ be a graph and $S, T \subseteq V$ with $\kappa(S, T) \geq d+1$. Choose algebraically independent edgeweights $c_{ij} > 0$. Map the nodes of S into $\mathbb{R}^d$ in general position. Then the rubber band extension of this map satisfies* $\mathrm{rk}(\mathbf{x}_T) = d+1$.

If the edgeweights are chosen randomly, independently and uniformly from $[0, 1]$, then the algebraic independence condition is satisfied with probability 1.

Proof. The proof will consist of two steps: first, we show that there is *some* choice of the edgeweights for which the conclusion holds; then we use this to prove that the conclusion holds for every generic choice of edge-weights.

For the first step, we use that by Menger's Theorem, there are $d+1$ disjoint paths P_i connecting a node $i \in S$ with a node $i' \in T$. The idea is to make the rubber bands on these paths very strong (while keeping the strength of the other edges fixed). Then these paths will pull each node i' very close to i. Since the positions of the nodes $i \in S$ are affine independent, so will be the positions of the nodes i' (Figure 1.3).

To make this precise, let D be the diameter of the set $\{\mathbf{x}_i : i \in S\}$ and let $E' = \cup_{i \in S} E(P_i)$. Fix any $R > 0$, and define a weighting $w = w_R$ of the edges by

$$w_e = \begin{cases} R, & \text{if } e \in E', \\ 1, & \text{otherwise.} \end{cases}$$

Let $\mathbf{x}$ be the rubber band extension of the given mapping of the nodes of S with these strengths.

Recall that f minimizes the potential $\mathcal{E}_w$ over all representations of G with the given nodes nailed. Let $\mathbf{y}$ be the representation with $\mathbf{y}_j = \mathbf{x}_i$ if $j \in P_i$ (in particular, $\mathbf{y}_i = \mathbf{x}_i$ if $i \in S$); for any node j not on any P_i, let $\mathbf{y}_j$ be any point in $\mathrm{conv}\{\mathbf{x}_i : i \in S\}$. In the representation $\mathbf{y}$ the edges with strength R have 0 length, and so

$$\mathcal{E}_w(\mathbf{x}) \leq \mathcal{E}_w(y) \leq D^2|E|.$$

On the other hand,

$$\mathcal{E}_w(\mathbf{x}) \geq \sum_{uv \in E'} R|\mathbf{x}_u - \mathbf{x}_v|^2.$$

By the triangle inequality and Cauchy-Schwartz Inequality we get that

$$\begin{aligned}|\mathbf{x}_i - \mathbf{x}_{i'}| &\leq \sum_{uv \in E(P_i)} |\mathbf{x}_u - \mathbf{x}_v| \leq \Big(|E(P_i)| \sum_{uv \in E(P_i)} |\mathbf{x}_u - \mathbf{x}_v|^2\Big)^{1/2} \\ &\leq \Big(\frac{n}{R}\mathcal{E}_w(\mathbf{x})\Big)^{1/2} \leq \frac{\sqrt{n|E|}D}{\sqrt{R}}.\end{aligned}$$

Thus $\mathbf{x}_{i'} \to \mathbf{x}_i$ if $R \to \infty$. Since $\{\mathbf{x}_i : \ i \in S\}$ are affine independent, this implies that $\{\mathbf{x}_{i'} : \ i \in S\} = \{\mathbf{x}_j : \ j \in T\}$ are affine independent.

This completes the first step. Now we argue that this holds for all generic edge-weightings. To prove this, we only need some general considerations. The embedding minimizing the energy is unique, and so it can be computed from the equilibrium equations (say, by Cramer's Rule). What is important from this is that the vectors $\mathbf{x}_i$ can be expressed as rational functions of the edgeweights. Furthermore, the value $\det((1+\mathbf{x}_{t_i}^{\mathsf{T}}\mathbf{x}_{t_j})_{i,j=0}^{d})$ is a polynomial in the coordinates of the $\mathbf{x}_i$, and so it is also a rational function of the edgeweights. We know that this rational function is not identically 0; hence it follows that it is nonzero for every algebraically independent substitution. □

1.3.3 Rubber bands and connectivity testing

Rubber band representations yield a (randomized) graph connectivity algorithm with good running time (Linial, Lovász and Wigderson [2]); however, a number of algorithmic ideas are needed to make it work, which we describe in their simplest form. We refer to the paper for a more thorough analysis.

Connectivity between given sets. We start with describing a test checking whether or not two given k-tuples S and T of nodes are connected by k node-disjoint paths. For this, we assign random weights c_{ij} to the edges, and map the nodes in S into $\mathbb{R}^{k-1}$ in general position. Then we compute the rubber band representation extending this map, and check whether the points representing T are in general position.

This procedure requires solving a system of linear equations, which is easy, but the system is quite large, and it depends on the choice of S and T. With a little work, we can save quite a lot.

Let $G = (V, E)$ be a connected graph on $V = [n]$ and $S \subseteq V$, say $S = [k]$. Let $d = k-1$. Given a map $\mathbf{x} : \ S \to \mathbb{R}^d$, we can compute its rubber band extension by solving the system of linear equations

$$\sum_{j \in N(i)} c_{ij}(\mathbf{x}_i - \mathbf{x}_j) = 0 \qquad (i \in V \setminus S). \tag{1.1}$$

This system has $(n-k)d$ unknowns and the same number of equations, and we know that it has a unique solution, since this is where the gradient of a strictly convex function (which tends to ∞ at ∞) vanishes.

At the first sight, solving (1.1) takes inverting an $(n-k)d \times (n-k)d$ matrix. However, we can immediately see that the coordinates can be computed independently, and since they satisfy the same equations except for the right hand side, it suffices to invert the matrix of the system once.

Below, we will face the task of computing several rubber band representations of the *same* graph, changing only the nailed set S. Can we make use of some of the computation done for one of these representations when computing the others?

The answer is yes. First, we do something which seems to make things worse: We create new "equilibrium" equations for the nailed nodes, introducing new variables for the forces that act on the nails. Let $\mathbf{y}_i = \sum_{j \in N(i)} c_{ij}(\mathbf{x}_j - \mathbf{x}_i)$ be the force acting on node i in the equilibrium position. Let X be the $V \times d$ matrix in which row i is $\mathbf{x}_i^\mathsf{T}$, and let Y be the $V \times d$ matrix in which row i is $\mathbf{y}_i^\mathsf{T}$. Let L_c be the symmetric $V \times V$ matrix

$$(L_c)_{ij} = \begin{cases} -c_{ij} & \text{if } ij \in E, \\ \sum_{k \in N(i)} c_{ik}, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

Then we can write (1.1) as $L_c X = Y$. To simplify matters a little, let us assume that the center of gravity of the representing nodes is 0, so that $JX = 0$. We clearly also have $JY = 0$ and $Y^\mathsf{T}\mathbf{e}_i = \mathbf{y}_i = 0$ for $i \notin S$.

Now we use the same trick as we did for harmonic functions: L_c is positive semidefinite, and "almost" invertible in the sense that the only vector in its nullspace is $\mathbb{1}$. Hence $L_c + J$ is invertible; let $M = (L_c + J)^{-1}$. Since $(L_c + J)J = J^2 = nJ$, it follows that $MJ = \frac{1}{n}J$. Furthermore, from $M(L_c + J) = I$ it follows that $ML_c = I - MJ = I - \frac{1}{n}J$. Hence $X = \big(I - \frac{1}{n}J\big)X = MY$.

It is not clear that we are nearer our goal, since how do we know the matrix Y (i.e., the forces acting on the nails)? But the trick is this: we can prescribe these forces arbitrarily, as long as $\sum_{i \in S} \mathbf{y}_i = 0$ and $\mathbf{y}_j = 0$ for $j \notin S$ are satisfied. Then the rows of $X = MY$ will give some position for each node, for which $L_c X = L_c MY = \big(I - \frac{1}{n}J\big)Y = Y$. So in particular the nodes in $V \setminus S$ will be in equilibrium, so we can just nail the nodes of S. If Y has rank d, then so does X, which implies that the vectors representing S will span the space $\mathbb{R}^d$. Since their center of gravity is 0, it follows that they are not on one hyperplane. We can apply an affine transformation to move the points $\mathbf{x}_i$ $(i \in S)$ to any other affine independent position if we wish, but this is irrelevant: what we want is to check whether the nodes of T are in

general position, and this is not altered by any affine transformation. A convenient choice is

$$Y = \begin{pmatrix} -1 \ldots -1 \\ I \\ 0 \end{pmatrix}$$

To sum up, to check whether the graph is k-connected between S and T (where $S, T \subseteq V$, $|S| = |T| = k$), we select random weights c_{ij} for the edges, compute the matrix $X = (L_c + J)^{-1}Y$, and check whether the rows with indices from T are affine independent.

Connectivity between all pairs. If we want to apply the previous algorithm for connectivity testing, it seems that we have to apply it for all pairs of k-sets. Even though we can use the same edgeweights and we only need to invert $L_c + J$ once, we have to compute $X = (L_c + J)^{-1}Y$ for potentially exponentially many different sets S, and then we have to test whether the nodes of T are represented in general position for exponentially many sets T. The following lemma shows how to get around this.

Lemma 1.3.6 *For every vertex $v \in V$ we select an arbitrary k-subset $S(v)$ of $N(v)$. Then G is k-connected iff $S(u)$ and $S(v)$ are connected by k node-disjoint paths for every u and v.*

Proof. The "only if" part follows from the well-known property of k-connected graphs that any two k-subsets are are connected by k node-disjoint paths. The "if" part follows from the observation that if $S(u)$ and $S(v)$ are connected by k node-disjoint paths, then u and v are connected by k openly disjoint paths. □

Thus the subroutine in the first part needs be called only $O(n^2)$ times. (We do not even have to check this for every pair (u, v), and further savings can be achieved through randomization, using Exercises 1.3.10 and 1.3.11.)

Numerical issues. The computation of the rubber band representation requires solving a system of linear equations. We have seen (see Figure 1.4) that for a graph with n nodes, the positions of the nodes can get exponentially close in a rubber band representation, which means that we might have to compute with exponentially small numbers (in n), which means that we have to compute with linearly many digits, which gives an extra factor of n in the running time. For computational purposes it makes sense to solve the system in a finite field rather than in $\mathbb{R}$. Of course, this "modular" embedding has no physical or geometrical meaning any more, but the algebraic structure remains!

Let $G = (V, E)$ be a graph and let $S \subseteq V$, $|S| = k$, and $d = k - 1$. Let p be a prime and $c_e \in \mathbb{F}_p$ for $e \in E$. A *modular rubber band representation* of G (with respect to S, p and c) is defined as an assignment $i \mapsto \mathbf{x}_i \in \mathbb{F}_p^d$ satisfying

$$\sum_{j \in N(i)} c_{ij}(\mathbf{x}_i - \mathbf{x}_j) = 0 \qquad (\forall i \in V \setminus S). \tag{1.2}$$

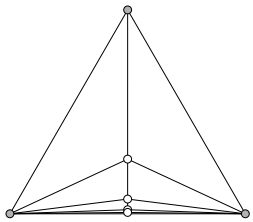

Figure 1.4: The rubber band embedding gives nodes that are exponentially close.

This is formally the same equation as for real rubber bands, but we work over $\mathbb{F}_p$, so no notion of convexity can be used. In particular, we cannot be sure that the system has a solution. But things work if the prime is chosen at random.

Lemma 1.3.7 *Let $N > 0$ be an integer. Choose uniformly a random a prime $p < N$ and random weights $c_e \in \mathbb{F}_p$ $(e \in E)$.*

(a) *With probability at least $1 - n^2/N$, there is a modular rubber band representation of G (with respect to S, p and c), such that the vectors $\mathbf{x}_i$ $(i \in S)$ are affine independent. This representation is unique up to affine transformations of $\mathbb{F}_p^d$.*

(b) *Let $T \subseteq V$, $|T| = k$. Then with probability at least $1 - n^2/N$, the representation $\mathbf{x}_c$ in* (a) *satisfies* $\mathrm{rk}(\{\mathbf{x}_i : i \in T\}) = \kappa(S, T)$.

Proof. The determinant of the system (1.2) is a polynomial of degree at most n^2 of the weights c_{ij}. The Schwartz–Zippel Lemma [4, 6] gives a bound on the probability that this determinant vanishes, which gives (a). The proof of (b) is similar. □

We sum up the complexity of this algorithm without going into fine details. In particular, we restrict this short analysis to the case when $k < n/2$. We have to fix an appropriate N; $N = n^5$ will be fine. Then we pick a random prime $p < N$. Since N has $O(\log n)$ digits, the selection of p and every arithmetic operation in $\mathbb{F}_p$ can be done in polylogarithmic time, and we are going to ignore them. We have to generate $O(n^2)$ random weights for the edges. We have to invert (over $\mathbb{F}_p$) the matrix $L_c + J$, this takes $O(M(n))$ operations, where $M(n)$ is the cost of matrix multiplication (currently known to be $O(n^{2.3727})$, Vassilevska Williams [5]). Then we have to compute $(L_c + J)^{-1}Y$ for a polylogarithmic number of different matrices Y (using Exercise 1.3.11(b) below). For each Y, we have to check affine independence for one k-set in the neighborhood of every node, in $O(nM(k))$ time. Up to polylogarithmic factors, this takes $O(M(n) + nM(k))$ work.

Exercise 1.3.8 A *convex representation* of a graph $G = (V, E)$ (in dimension d, with boundary $S \subseteq V$) is an mapping of $V \to \mathbb{R}^d$ such that every node in $V \setminus S$ is in the relative interior of the convex hull of its neighbors.(a) The rubber band representation extending any map from $S \subseteq V$ to $\mathbb{R}^d$ is convex with boundary S. (b) Not every convex representation is constructible this way.

Exercise 1.3.9 A 1-dimensional convex representation with two boundary nodes s and t is called an $s-t$*-numbering*. (a) Prove that every 2-connected graph $G = (V, E)$ has an $s-t$-numbering for every $s, t \in V$, $s \neq t$. (b) Show that instead of the 2-connectivity of G, it suffices to assume that deleting any node, either the rest is connected or it has two components, one containing s and one containing t.

Exercise 1.3.10 Let G be any graph, and let H be a k-connected graph with $V(H) = V(G)$. Then G is k-connected if and only if u and v are connected by k openly disjoint paths in G for every edge $uv \in E(H)$.

Exercise 1.3.11 Let G be any graph and $k < n$. (a) For t pairs of nodes chosen randomly and independently, we test whether they are connected by k openly disjoint paths. Prove that if G is not k-connected, then this test discovers this with probability at least $1-\exp(-2(n-k)t/n^2)$. (b) For r nodes chosen randomly and independently, we test whether they are connected by k openly disjoint paths to every other node of the graph. Prove that if G is not k-connected, then this test discovers this with probability at least $1-((k-1)/n)^r$.

Exercise 1.3.12 Let $G = (V, E)$ be a graph, and let Z be a matrix obtained from L_G by replacing its nonzero entries by algebraically independent transcendentals. For $S, T \subseteq V(G)$, $|S| = |T| = k$, let $Z_{S,T}$ denote the matrix obtained from Z by deleting the rows corresponding to S and the columns corresponding to T. Then $\det(Z_{S,T}) \neq 0$ if and only if $\kappa(S, T) = k$.

1.4 Harmonic functions on graphs

The notion of a harmonic function is basic in analysis, usually defined as smooth functions $f : \mathbb{R}^d \to \mathbb{R}$ (perhaps defined only in some domain) satisfying the differential equation $\Delta f = 0$, where $\Delta = \sum_{i=1}^{d} \partial^2/\partial x_i^2$ is the Laplace operator. It is a basic fact that such functions can be characterized by the "mean value property" that their value at any point equals to their average on a ball around this point. Taking this second characterization as our starting point, we will introduce an analogous notion of harmonic functions defined on the nodes of a graph.

Harmonic functions play an important role in the study of random walks (after all, the averaging in the definition can be interpreted as expectation after one move). They also come up in the theory of electrical networks, and also in statics. This provides a connection between these fields, which can be exploited. In particular, various methods and results from the theory of electricity and statics, often motivated by physics, can be applied to provide results about random walks. From our point of view, their main applications will be rubber band representations and discrete analytic functions. In this section we develop this simple but useful theory.

1.4.1 Definition and uniqueness

Let $G = (V, E)$ be a connected simple graph and $S \subseteq V$. For a function $f : V \to \mathbb{R}$, we define its *defect* at node v by

$$(Lf)(v) = \sum_{u \in N(v)} (f(u) - f(v)). \tag{1.3}$$

The function f is called a *harmonic* at a node $v \in V$ if $(Lf)(v) = 0$. This equation can also be written in the form

$$\frac{1}{d_v} \sum_{u \in N(v)} f(u) = f(v). \tag{1.4}$$

A node where a function is not harmonic is called a *pole* of the function. It is useful to note that

$$\sum_{v \in V} (Lf)(v) = \mathbb{1}^{\mathsf{T}} Lf = f^{\mathsf{T}} L \mathbb{1} = 0, \tag{1.5}$$

so every function is harmonic "on the average".

We can extend this notion to multigraphs by replacing (1.4) by

$$\frac{1}{d_v} \sum_{\substack{e \in E \\ V(e) = \{u,v\} \ni v}} f(u) = f(v) \qquad \forall v \in V \setminus S. \tag{1.6}$$

Another way of writing this is

$$\sum_{u \in V} a_{uv}(f(u) - f(v)) = 0 \qquad \forall v \in V \setminus S, \tag{1.7}$$

where a_{uv} is the multiplicity of the edge uv. We could further generalize this by assigning arbitrary nonnegative weights β_{uv} to the edges of G ($uv \in E(G)$), and require the condition

$$\sum_{u \in V} \beta_{uv}(f(u) - f(v)) = 0 \qquad \forall v \in V \setminus S. \tag{1.8}$$

Every constant function is harmonic at each node. On the other hand,

Proposition 1.4.1 *Every non-constant function has at least two poles.*

Proof. Let S be the set where the function assumes its maximum, and let S' be the set of those nodes in S that are connected to any node outside S. Then every node in S' must be a pole, since in (1.4), every value $f(u)$ on the left had side is at most $f(v)$, and at least one is less, so the average is less than $f(v)$. Since the function is nonconstant, S is a nonempty proper subset of V, and since the graph is connected, S' is nonempty. So there is a pole where the function attains its maximum. Similarly, there is another pole where it attains its minimum. □

For any two nodes there will be a nonconstant harmonic function that is harmonic everywhere else. More generally, we have the following theorem.

Theorem 1.4.2 *For every nonempty set $S \subseteq V$ and every function $f_0 : S \to \mathbb{R}$ there is a unique function $f : V \to \mathbb{R}$ extending f_0 that is harmonic at each node of $V \setminus S$.*

We call this function f the *harmonic extension* of f_0. Note that if $|S| = 1$, then the harmonic extension is a constant function (and so it is also harmonic at S, and it does not contradict Proposition 1.4.1). If $S = \{a, b\}$, then a function that is harmonic outside S is uniquely determined by two of its values, $f(a)$ and $f(b)$. Scaling the function by a real number and translating by a constant preserves harmonicity at each node, so if we know the function f with (say) $f(a) = 0$ and $f(b) = 1$, harmonic at $v \neq a, b$, then $g(v) = A + (B - A)f(v)$ describes the harmonic extension with $g(a) = A$, $g(b) = B$. In this case equation (1.4) is equivalent to saying that $F_{ij} = f(j) - f(i)$ defines a flow from a to b.

The uniqueness of the harmonic extension is easy by the argument in the proof of Proposition 1.4.1. Suppose that f and f' are two harmonic extensions of f_0. Then $g = f - f'$ is harmonic on $V \setminus S$, and satisfies $g(v) = 0$ at each $v \in S$. If g is the identically 0 function, then $f = f'$ as claimed. Else, either is minimum or its maximum is different from 0. But we have seen that both these sets contain at least one pole, which is a contradiction.

1.4.2 Constructing harmonic functions

There are several ways to construct a harmonic function with given poles, and we describe four of them. This abundance of proof makes sense, because each of these constructions illustrates an application area of harmonic functions.

Linear algebra

First we construct a harmonic function with two given poles a and b. Let $\mathbb{1}_a : V \to \mathbb{R}$ denote the function which is 1 on a and 0 everywhere else. Consider the equation

$$Lf = \mathbb{1}_b - \mathbb{1}_a. \tag{1.9}$$

(where L is the Laplacian of the graph). Every function satisfying this equation is harmonic outside $\{a, b\}$.

The Laplacian L is not quite invertible, but its nullspace is only one-dimensional, spanned by the vector $\mathbb{1} = (1, \ldots, 1)^\mathsf{T}$. So (1.9) determines f up to adding the same scalar to every entry. Hence we may look for a special solution satisfying $\mathbb{1}^\mathsf{T} f = 0$, or equivalently, $Jf = 0$ (where $J \in \mathbb{R}^{V \times V}$ denotes the all-1 matrix). The trick is to observe that $L + J$ is invertible, and

$$(L + J)f = Lf = \mathbb{1}_b - \mathbb{1}_a,$$

and so we can express f as

$$f = (L + J)^{-1}(\mathbb{1}_b - \mathbb{1}_a). \tag{1.10}$$

All other solutions of (1.9) can be obtained from f by adding the same scalar to every entry; all other functions harmonic outside $\{a, b\}$ can be obtained from these by multiplying by a scalar.

Second, suppose that $|S| \geq 3$, let $a, b \in S$, and let $g_{a,b}$ denote the function which is harmonic outside $\{a, b\}$ and satisfies $g_{a,b}(a) = 0$ and $g_{a,b}(b) = 1$. We then have

$$Lg_{a,b}(x) \begin{cases} \neq 0, & \text{if } x \in \{a, b\}, \\ = 0, & \text{otherwise.} \end{cases}$$

Note that every linear combination of the functions $g_{a,b}$ is harmonic outside S.

Let $g'_{a,b}$ denote the restriction of $g_{a,b}$ to S. Fix $a \in S$. We claim that the functions $g'_{a,b}$ ($b \in S \setminus \{a\}$) are linearly independent. Indeed, a relation $\sum_b \alpha_b g'_{a,b} = 0$ implies that $\sum_b \alpha_b g_{a,b} = 0$ (by the uniqueness of the harmonic extension), which in turn implies that $\sum_b \alpha_b L g_{a,b} = 0$. The value of this last function at $b \in S \setminus \{a\}$ is $\alpha_b L g_{a,b}(b)$, where $Lg_{a,b}(b) \neq 0$, and so $\alpha_b = 0$.

It follows that the functions $g'_{a,b}$ generate the $(|S|-1)$-dimensional space of all functions $h \in \mathbb{R}^S$ with $h(a) = 0$. This means that $f_0 - f_0(a)$ is a linear combination of functions $g'_{a,b}$ and a constant function. The corresponding linear combination of the $g_{a,b}$ the harmonic extension of $f_0 - f_0(a)$. Adding the constant $f_0(a)$ at every node, we get the harmonic extension of f_0.

Random walks

Let $G = (V, E)$ be a connected graph and $v \in V$. Start a random walk $W_v = (w^0, w^1, \dots)$ at $w^0 = v$. For any event A expressible in terms of W_v, we denote by $\mathsf{P}_v(A)$ its probability. Expectations $\mathsf{E}_v(X)$ are defined similarly for random variables X depending on the walk W_v. (One needs some facts from the theory of random walks, for example, that a random walk hits every node with probability 1.)

Fixing a nonempty set $S \subseteq V$, let $f_a(v)$ denote the probability that W_v hits $a \in S$ before it hits any other element of S. Then trivially

$$f_a(v) = \begin{cases} 1, & \text{if } v = a, \\ 0, & \text{if } v \in S \setminus \{a\}. \end{cases}$$

For $v \in V \setminus S$, we have

$$f_a(v) = \mathsf{P}_v(W_v \text{ hits } S \text{ at } a) = \sum_{u \in N(v)} \mathsf{P}_v(w^1 = u)\mathsf{P}_u(W_u \text{ hits } S \text{ at } a) = \frac{1}{d(v)} \sum_{u \in N(v)} f_a(u).$$

This shows that f_a is harmonic at all nodes outside S. Since every function on S can be written as linear combinations of functions $f_a|_S$, it follows that every function on S has a harmonic extension.

This harmonic extension can be given a more direct interpretation. Given a function $f_0 : S \to \mathbb{R}$, we define $f(v)$ for $v \in V$ as the expectation of $f_0(\mathbf{a}_v)$, where $\mathbf{a}_v$ is the (random)

node where a random walk W_v first hits S. Clearly $f(v) = f_0(v)$ for $v \in S$, and it is easy to check (by a similar computation as above) that $f(v)$ is harmonic at $V \setminus S$.

Electrical networks

Consider the graph G as an electrical network, where each edge represents a unit resistance. Assume that an electric current is flowing through G, entering at a and leaving at b. Let $f(v)$ be the potential of node v. Then f is a harmonic except at a and b. Indeed, the current through an edge uv is $f(u)-f(v)$ by Ohm's Law, and hence by Kirchhoff's Current Law, $\sum_{j\in N(i)}(f(j)-f(i)) = 0$.

More generally, given a set $S \subseteq V$ and a function $f_0 : S \to \mathbb{R}$, let a current flow through the network while keeping each $v \in S$ at potential $f_0(v)$. Then the potential $f(v)$ defined for all nodes v is an extension of f_0. The same argument as above implies that f is harmonic at every node $v \in V \setminus S$.

Rubber bands

Consider the edges of the graph G as ideal rubber bands (or springs) with unit Hooke constant (i.e., it takes h units of force to stretch them to length h). Grab the nodes a and b and stretch the graph so that their distance is 1. Let the structure find its equilibrium, and let $f(v)$ be the distance of node v from a. Since these nodes are at equilibrium, the function f is harmonic for $v \neq a, b$, and satisfies $f(a) = 0$ and $f(b) = 1$. To see the last fact, note that the edge uv pulls v with force $f(u)-f(v)$ (the force is positive if it pulls in the positive direction), and so (1.4) is just the condition of the equilibrium.

More generally, if we have a set $S \subseteq V$ and we fix the positions of each node $v \in S$ at a given point $f_0(v)$ of the real line, and let the remaining nodes find their equilibrium, then the position $f(v)$ of a node will be a function that extends f_0 (trivially), and that is harmonic at each $v \notin S$.

We'll come back to this construction repeatedly, extending it to higher dimensions. We'll prove that the equilibrium exists and is unique without any reference to physics.

1.4.3 Relating different models

A consequence of the uniqueness property in Theorem 1.4.2 is that the harmonic functions constructed in sections 1.4.2, 1.4.2, 1.4.2 and 1.4.2 are the same. As an application of this idea, we show the following interesting identities (see Nash-Williams [3], Chandra at al. [1]).

Considering the graph G as an electrical network, where every edge has resistance 1, let $R(s,t)$ denote the effective resistance between nodes s and t. Considering the graph G as a rubber band structure in equilibrium, with two nodes s and t nailed down at 1 and 0, let F_{ab} denote the force pulling the nails. Doing a random walk on the graph, let $\mathsf{comm}(a,b)$ denote

the commute time between nodes a and b (i.e., the expected time it takes to start at a, walk until you first hit b, and then walk until you first hit a again).

Theorem 1.4.3 *The effective resistance $R(a,b)$ between two nodes a, b of a graph, the force F_{ab} needed to pull them distance 1 apart, and the commute time $\mathsf{comm}(a,b)$ of the random walk between them, are related by the equations*

$$R(a,b) = \frac{1}{F_{ab}} = \frac{\mathsf{comm}(a,b)}{2m}.$$

Proof. Consider the function $f \in \mathbb{R}^V$ satisfying $f(a) = 0$, $f(b) = 1$ and harmonic for $v \neq a, b$ (we know that this is uniquely determined). We are going to express the quantities in the formula by $(Lf)(a) = \sum_{v \in N(a)} f(v)$ (this will prove the theorem):

$$R(a,b) = \frac{1}{(Lf)(a)}, \qquad F_{ab} = (Lf)(a), \qquad \mathsf{comm}(a,b) = \frac{2m}{(Lf)(a)}. \tag{1.11}$$

By the discussion in section 1.4.2, $f(v)$ is equal to the potential of v if we fix the potential of a and b. The current through the network is $(Lf)(a)$. So the effective resistance is $R(a,b) = 1/(Lf)(a)$.

The equation $F_{ab} = (Lf)(a)$ is easily derived from Hooke's Law.

Finally, we know that $f(u)$ is the probability that a random walk starting at u visits a before b. Hence $p = (1/d(a)) \sum_{u \in N(a)} f(u) = (Lf)(a)/d(a)$ is the probability that a random walk starting at a hits b before returning to a.

Let $\mathbf{T}$ be the first time when a random walk starting at a returns to a and $\mathbf{S}$, the first time when it returns to a after visiting b. We know from the theory of random walks that $E(\mathbf{T}) = 2m/d(a)$ and by definition, $E(\mathbf{S}) = \mathsf{comm}(a,b)$. Clearly $\mathbf{T} \leq \mathbf{S}$ and $\mathbf{T} = \mathbf{S}$ means that the random walk starting at a hits b before returning to a. Thus $\mathsf{P}(\mathbf{T} = \mathbf{S}) = p$. If $\mathbf{T} < \mathbf{S}$, then after the first $\mathbf{T}$ steps, we have to walk from a until we reach b and then return to b. Hence $E(\mathbf{S} - \mathbf{T}) = (1-p)E(\mathbf{S}) = (1-p)\mathsf{comm}(a,b)$, and so

$$\frac{2m}{d(a)} = \mathsf{E}(\mathbf{T}) = \mathsf{E}(\mathbf{S}) - \mathsf{E}(\mathbf{S} - \mathbf{T}) = \mathsf{comm}(a,b) - (1-p)\mathsf{comm}(a,b) = \frac{(Lf)(a)}{d(a)}\mathsf{comm}(a,b).$$

Simplifying, we get 1.11. □

For the rubber band model, imagine that we slowly stretch the graph until nodes a and b will be at distance 1. When they are at distance t, the force pulling our hands is tF_{ab}, and hence the energy we have to spend is

$$\int_0^1 tF_{ab}\, dt = \frac{1}{2}F_{ab}.$$

By conservation of energy, we get the identity

$$\sum_{ij \in E} (f(i) - f(j))^2 = F_{ab}. \tag{1.12}$$

Using the "topological formulas" from the theory of electrical networks for the resistance, we get a further characterization of these quantities, which we state without proof. Let G' denote the graph obtained from G by identifying a and b, and let $\mathcal{T}(G)$ denote the number of spanning trees of G. Then

$$R(a,b) = \frac{\mathcal{T}(G)}{\mathcal{T}(G')}. \tag{1.13}$$

Another application of this equivalence between our three models is the following fact.

Theorem 1.4.4 *Let $G = (V, E)$ be a connected graph, $a, b \in V$, and let us connect two nodes by a new edge.*

(a) *The ratio $\mathsf{comm}(a,b)/|E|$ does not increase.*

(b) (Raleigh's Theorem) *The effective resistance $R(a,b)$ does not increase.*

(c) *If nodes a and b nailed down at 0 and 1, the force pulling the nails in the equilibrium does not decrease.*

Proof. The three statements are equivalent by Theorem 1.4.3; we prove (c). By (1.12), it suffices to prove that the equilibrium energy does not decrease. Consider the equilibrium position of G', and delete the new edge. The contribution of this edge to the energy was nonnegative, so the total energy decreases. The current position of the nodes may not be in equilibrium; but the equilibrium position minimizes the energy, so when they move to the equilibrium of G', the energy further decreases. □

Exercise 1.4.5 Show how L^{-1} can be computed by inverting the nonsingular matrix $L + J$.

Exercise 1.4.6 Let a, b be two nodes of a connected graph G. We define the *hitting time* $H(a,b)$ as the expected number of steps before a random walk, started at a, will reach b. Consider the graph as a rubber band structure as above, but also attach a weight of $d(v)$ to each node v. Nail the node b to the wall and let it find its equilibrium. Prove that a will be at distance $H(a,b)$ below b.

Exercise 1.4.7 Show by an example that the commute time $\mathsf{comm}(a,b)$ can increase when we add an edge to the graph.

Bibliography

[1] A.K. Chandra, P. Raghavan, W.L. Ruzzo, R. Smolensky and P. Tiwari: The electrical resistance of a graph captures its commute and cover times, *Proc. 21st ACM STOC* (1989), 574–586.

[2] N. Linial, L. Lovász, A. Wigderson: Rubber bands, convex embeddings, and graph connectivity, *Combinatorica* **8** (1988) 91-102.

[3] C. St.J. A. Nash-Williams, Random walks and electric currents in networks, *Proc. Cambridge Phil. Soc.* **55** (1959), 181–194.

[4] J. Schwartz: Fast probabilistic algorithms for verification of polynomial identities, *Journal of the ACM* **27** (1980), 701–717.

[5] V. Vassilevska Williams: Multiplying matrices faster than Coppersmith-Winograd,

[6] R. Zippel: Probabilistic algorithms for sparse polynomials, in: *Symbolic and Algebraic Computation*, Lecture Notes in Computer Science **72** (1979), 216–226.