THE INVERSE PROBLEM FOR DIRECTED CURRENT ELECTRICAL NETWORKS

JOEL NISHIMURA

Abstract. This paper investigates the inverse problem for the directed current networks defined by Orion Bawdon [1]. Using a conjecture about recoverability in the undirected case it will be shown that when each edge has conductivities in both directions and the graph is recoverable in the non-directed case then the directed electrical network is recoverable. Additionally, a method of solving graphs which have different forward conductivities than backward conductivities on each edge will be provided.

1. Introduction

It should be noted that the type of directed electrical network discussed in this paper is significantly different than the type of directed graphs described in previous papers in the UW Math REU. Future references to "directed" should be interpreted as referring to directed current electrical networks. Informally, in this type of electrical network conductances are associated with a direction where the conductance is only used if current agrees with that direction. The difference between these networks and the undirected case is essentially that in the directed current case the Dirichlet to Neumann (response) map is piecewise linear while in the undirected case the same map is linear.

For the remainder of the paper we will work on an electrical network $\Gamma = (G, \gamma)$ with graph G and conductivities γ. The graph G can be divided into vertices and directed edges that connect vertices. Let the vertices V of G be partitioned into $V = \partial V \cup \text{int } V$, where $\partial V \neq \emptyset$. Let there be n boundary nodes ∂V.

In the regular case the inverse problem could be thought of as taking a complete graph, or K graph, described by a single matrix and interiorizing it. However, not only is there no single matrix description there is also no K graph that can in general summerize the response of directed current graphs. Indeed there is not even a Y-Δ transformation.

Without a response matrix we now redefine the inverse problem on directed networks as a question of inputs and outputs.

Definition 1.1. *(Inverse Problem on Directed Current)* The inverse problem is to determine the conductances on a known directed graph G using the information gained by taking one parameter families of boundary potentials and measuring the resulting boundary currents.

This formulation of the inverse problem is similar to being given ideal power supplies and multimeters, taking measurements and then trying to determine the conductances inside.

Date: August 15, 2005.

This paper will focus on a particular class of directed current electrical networks one on a bidirected graph.

Definition 1.2. *(Bidirected Graph)* A graph is bidirected if between any two nodes there are either exactly zero edges or two edges, one in each direction. A bidirected electrical network or bidirected network is a bidirected graph with conductivities. A bidirected edge is a pair of edges that share vertices but have opposite directions.

From this definition it becomes clear that regular electrical networks are the subset of bidirected networks where the conductivities between two vertices are the same in both directions.

Definition 1.3. *(Friendly Graph)* A graph is friendly if it could be constructed by taking a recoverable regular graph and replacing every edge with a bidirected edge such that the final graph is bidirected.

It will be shown if a conjecture is assumed that the a bidirected graph is recoverable if and only if it is friendly. Each bidirected network will have multiple undirected networks associated with it and the recovery algorithms for these undirected networks will be utilized to recover the conductances for the bidirected network. However, before that can be directly shown it is important to look at the properties that the response of directed current networks have.

2. The Response of Directed Networks

In the undirected case since the entire response is summed up in a single matrix it is not necessary to investigate the space of different boundary potentials. The directed case will be shown to be more complicated. Since there are n different boundary nodes the space of different inputs is n dimensional, the voltage at each boundary node can be changed independently of each other. Additionally the proof of existence and uniqueness for the Dirichlet problem provides that all possible inputs provide a solution and each input provides only one solution. We can thus view all possible inputs of our inverse problem existing in n-space, or voltage space. This space can be reduced by taking into consideration the properties of electrical networks. Since adding the same constants to each boundary node has no effect on current the space can be reduced to an $n-1$ hyperplane perpendicular to the vector [1,...1]. Furthermore since scaling boundary voltages also scales currents this $n-1$-space can be described as a $n-2$ sphere centered at the origin and a scaling factor. Thus the possible responses on a graph with 3 boundary nodes could be described by inputs placed on a circle perpendicular to the vector [1,...1] and a scaling factor. This spatial view of different inputs is important because this space is divided into areas of different linear behavior.

Because a voltage input produces only one solution each point in the voltage space has a definitive current pattern where current pattern is defined as:

Definition 2.1. *(Current Pattern)* A current pattern is the association of an orientation to each bidirected edge. A network with boundary voltages V_b is said to have a current pattern P if the orientation of the currents in V_b does not disagree with the orientations in P. (Note: when the current on an edge is zero it has no orientation.)

Notice that with this definition some points in voltage space may have multiple current patterns. This is allowed to occur when an edge in the graph carries no

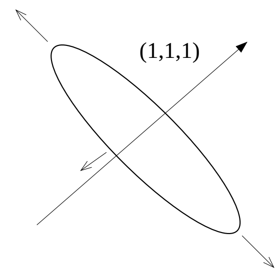

FIGURE 2.1. The entire response for 3-space can be described by a ring perpendicular to the line $(1, 1, 1)$ and a scaling factor.

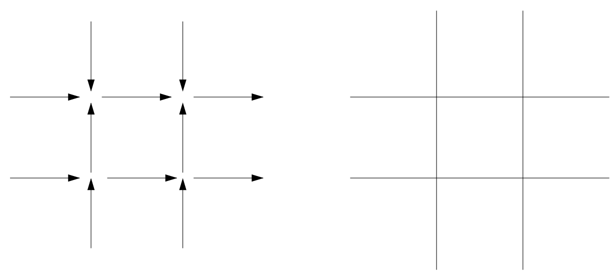

FIGURE 2.2. This is one of the many possible current patterns on a 3 by 3 lattice.

current, the valid current patterns will agree with each other on all edges that carry current but are allowed to disagree on edges that carry no current. The solution to the Dirichlet problem for directed networks [1] also implies that every point has a current pattern and the uniqueness of the Dirichlet problem for bidirected networks [1] implies that at each point where no edge has zero current the graph has only one current pattern.

For each current pattern the bidirected graph can be described by an undirected network; simply create an undirected network using the conductances from the bidirected network that agree with the orientations in the current pattern. This implies that the response for each current pattern is linear.

Definition 2.2. *(Domain of Linear Behavior)* A domain of linear behavior is the locus of different boundary potentials that result in the same current pattern. These will be referred to as simply "domains" for the remainder of the paper.

Since all boundary voltages produce a current pattern then all boundary voltages lie in at least one domain. Furthermore, domains can only overlap when an edge has zero current on it by the way that current pattern was defined.

Domains have the property that linear positive combinations of points in a domain remain in that domain. Since every point in a domain has the same direction of current on each edge positive linear combinations of these points will, by superposition of undirected graphs produce another point with the same current pattern and that point will thus lie in the same domain.

This leads rather directly to the conclusion that domains are convex and that the Dirichlet to Neumann map is continuous. Also note that the borders between different domains are by definition the borders between different current patterns and because the response in each domain is linear and thus continuous these borders between domains must have the property that an edge has zero current.

Definition 2.3. *(Zero-Surface)* Zero-surface Z_i is the locus of points on which the current on edge i is zero.

Every edge in a friendly graph has a zero-surface because all edges have zero current when the boundary voltages are all equal.

Lemma 2.1. *Zero-Surfaces of friendly graphs are of dimension $n-1$*

Proof. Requiring an edge to have zero current is the same as requiring that that edge's vertices have equal voltages, $V_i = V_j$. This additional restraint equation is either independent or dependent with the original Kirchhoff requirements. If this restraint is dependent then the edge is always zero which contradicts recoverability in the undirected case and thus the graph isn't friendly. If the restraint is independent then adding the restraint will reduce the dimension to $n-1$. □

If one is given a Kirchhoff matrix for a undirected graph it is relatively straight forward to calculate the various zero-surfaces. Using the block format of the Kirchhoff matrix used in [2] the voltage in the interior nodes, V_I can be written as $-C^{-1}B^tV_b = 0$. Setting two neighboring interior nodes, V_i and V_j, equal to each other is the same as requiring there be zero current on the edge that connects them and yields the equation, $[-C^{-1}B^t]_iV_b = [-C^{(}-1)B^t]_jV_b$. This is simply a linear combination of the boundary voltages. Because the zero surfaces have the form of a linear combination of boundary voltages then zero-surfaces either always coincide or transversly intersect.

Calculating the zero-surfaces for a bidirectional graph is slightly more complicated as a Kirchhoff matrix can only be used to calculate the zero-surfaces which have the same current patterns. Thus creating an entire zero-surface involves patching together many different, smaller zero-surfaces.

Our next immediate goal will be to show that while two zero surfaces can intersect they can never coincide. While we desire this result on a friendly graph it will suffice to show it for a undirected recoverable graph. Since every point in the voltage space can be described by a linear region, showing that in these linear regions zero-surfaces can't coincide will imply that this is also true in the friendly case. Unfortunetly at this time a complete argument for this does not exist and we must conjecture it.

Conjecture 2.2. *Two zero-surfaces can't coincide on an undirected recoverable graph.*

There are a number of good reasons to believe that this must true. A more physical argument follows.

Suppose one has an undirected recoverable graph on which the two zero surfaces, Z_i and Z_j coincide. Let us place a voltage of one at boundary node ∂V_1 and measure the resulting current on edges i and j. The voltage at ∂V_1 will either create a current on both edges in a proportion c or will not generate any current on these edges. Let us continue this procedure for each boundary node. Suppose that there where two boundary nodes that created currents on i and j in a proportion not equal to c, then by some combination of these two boundary potentials current on one edge could be made to be zero on one and not the other contradicting the assumption

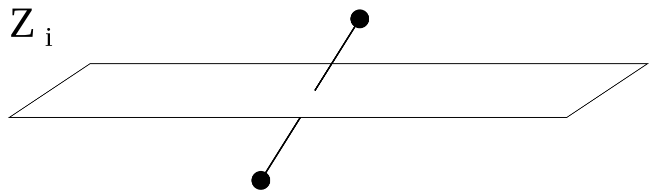

FIGURE 2.3. Here a line connects two points on opposite sides of the surface Z_i such that the midpoint of that line lies on Z_i.

that Z_i and Z_j coincide. Thus if Z_i and Z_j coincide then the voltage difference on i must always be proportional to the difference on j.

This, almost certainly raises an issue with recoverability. If the voltage difference on one edge is always exactly proportional to the voltage difference on another edge then how can the effect that one of those edges ever be differentiated from the effect of the other edge. The two edges effectively act like one component. Indeed this is exactly the problem that is encountered when trying to recover a parallel or series edge, the net effect can be determined but it is impossible to distinguish the effect that one edge has from the other.

Lemma 2.3. *Transversely crossing a single zero-surface Z_i causes the current on edge i to flip.*

Proof. Let us restrict our attention to a single crossing of the surface Z_i where Z_i does not intersect any other zero surface. Assume that on both sides of this single crossing the currents on edge i are in the same direction. Then the opposite sides of the crossing have the same current patterns and thus these and the zero-surface can be expressed by the same linear system. Now choose two points in the opposite domains such that the midpoint of the line connecting these points lies on the zero-surface. The sum of these points lies in the same domain as does the positive scaling of that sum. Notice though that the current of this average on edge i must be greater than zero as the two points used in that average both had currents on edge i that were in the same direction. Yet this nonzero current on edge i in Z_i is a contradiction. □

Theorem 2.4. *On a friendly graph by adjusting boundary voltages, current can flow on in either direction on any edge assuming conjecture 2.2.*

Proof. In a friendly graph every edge has an associated zero-surface of dimension $n-1$, and by conjecture none of these surfaces can coincide with any other surface. Thus each zero-surface can be individually transversely crossed implying that the current on that edge must flip. Thus any edge must express current flow in both directions in a friendly graph. □

3. The Recovery Method

Definition 3.1. *Boundary Voltage Partial Derivative* Let V_b be an n-dimensional vector of voltages at the boundary nodes and let $I_b(x)$ be defined as the n-dimensional vector of boundary currents produced by placing a voltage $\vec{x}$ at the boundary nodes. The positive derivative is $\partial(p,+)(V_b) = lim_{\epsilon->0}\frac{I_b(V_b+\epsilon_p)-I_b(V_b)}{\epsilon}$, where ϵ_p is equal to the n-dimensional vector with zeros in all entries except the pth entry, where there is a positive epsilon. The negative derivative is $\partial_(p,-)(V_b) = lim_{\epsilon->0}\frac{I_b(V_b-\epsilon_p)-I_b(V_b)}{\epsilon}$.

With this background it is now possible to go about providing the recovery method.

Theorem 3.1. *All friendly graphs are recoverable assuming conjecture 2.2.*

Proof. Assume that one is given a friendly graph. Randomly choose a point on the $n-2$ voltage sphere centered at the origin. We must now test to check if it is on a zero-surface. Take the positive and negative partial boundary voltage derivatives. The idea behind this test is if the point is on a zero-surface the partial derivative of the currents will be discontinuous as on one side there is a different linear system than on the other side. If $\partial_(p,+)(V_b) = -\partial_(p,-)(V_b)$ for all p then continue with the algorithm. If this is not true then there will be a p for which the derivative is discontinuous. If there is a p such that the derivative is discontinuous pick a new point along that direction a short distance away and test again. This will guarantee that the point has moved off one zero-surface, since there are a finite number of zero-surfaces eventually a point will be found that is in the interior of a region.

Once a point Q is found that does not lie on a zero-surface the core of the algorithm can be run. Recall that in the undirected case each row of the response matrix is essentially $\partial_(p,+)(V_b)$. Because the system is linear these derivatives are constant and can be linearly combined to create any desired boundary voltage. Since all of n-space is described by the same linear system then where these derivatives was computed was irrelevant. In this case though there are multiple linear domains, but like the undirected case the response for a domain can be described by the derivatives at a single point. Thus take the n derivatives at a the point Q and assemble the response matrix for that domain.

We now have a response matrix that corresponds to a graph that is recoverable in the undirected case, we recover that graph. This yields a Kirchhoff matrix K_0 valid for the domain that contains Q. While we have recovered a conductance for each edge in the graph we don't yet know which direction each edge is valid for. In order to find the resulting current pattern to associate with Q's domain we solve $KV = \Psi$. Combining the current pattern with the conductances we are able to specify which conductances are linked to which directions.

We now know the linear behavior in the region that contains Q, but we have yet to determine exactly the size and shape of that domain. However, the shape of the domain is stored in the Kirchoff matrix, because that can be used to solve for when edges are zero. Using the equation $[-C^(-1)B^t]_iV_b = [-C^(-1)B^t]_jV_b$ it is possible to plot all the zero-surfaces as that domain predicts (in fact only the parts of the zero-surfaces that border the domain are described by these equations). The domain is thus the convex region containing Q bounded by the union of the zero-surfaces. For the next point in the algorithm choose any point not in any domain yet recovered. Since there are finite domains the algorithm converges and because for any edge there are current patterns with opposite orientations on that edge all conductances can be recovered. □

Notice that if conjecture 2.2 is incorrect then not all edges might be able to be recovered. Thus, the recovery algorithm would create a graph that had the exact same response as the desired graph, but specified nothing for the conductance that was never used. Certainly all graphs that are recoverable in the undirected case and have the property that no two zero-surfaces coincide are recoverable in the bidirected case.

Corollary 3.1.1. *Nonfriendly graphs are not recoverable in general.*

Proof. Suppose one has a graph where each bidirectional edge has equal conductances in both directions. This graph is essentially the same as an undirected graph, yet since it is not friendly it is not recoverable. □

We now investigate a situation where by assuming that the bidirectional edges are unequal we are able to recover networks that are not recoverable in the undirected case.

4. Strictly Bidirectional Ring Trick

Definition 4.1. *Strictly Bidirectional* A strictly bidirectional graph is a bidirectional graph where the conductances on a bidirectional edge are not equal.

If it is assumed that the graph is strictly bidirectional then some additional classes of graphs become recoverable. This is because one can use the fact that the edges are strictly bidirectional as a test for the interior of the graph, in particular when the current on an edge flips direction the derivative of the boundary currents become discontinuous.

We now introduce the Strictly Bidirectional Ring Trick.

Let us begin with the delta spike graph which is a delta composed of interior nodes with one boundary spike attached to each interior node. Let us assign a boundary voltage of one at node one and zeros at boundary nodes two and three. From this pattern of boundary voltages we have three possible patterns of currents on the graph with the differences being on the five-six edge, there being no current, current from five to six or current from six to five. Notice that all the other edges have current directions that are determined, mainly that the current heads from the node one to nodes two and three. Also notice that on this graph no two edges can have zero current at the same time unless all the edges have zero current. This implies that the currents on this graph must change their direction one at a time since, by continuity, the currents must travel through zero before switching direction.

The fact that only one edge at a time can be zeroed out implies that a current is zeroed out if and only if in terms of the boundary voltages the boundary currents have a discontinuous derivative.

We will now see that by varying the voltage at node three there will be forced to be some point where the voltage on the five-six edge will be zero. We will now divide the three initial cases into three different cases. If the current is initially zero then the derivative of the boundary currents will be discontinuous at zero, clearly observable as the voltage at node three is increased and decreased.

Should the current be in the direction of five to six then by raising the voltage at node three eventually the current on the five-six must switch direction and decreasing the voltage at node three will not change the five-six edge before the current on edge two-five changes direction.

If the current is initially in the six-five direction then raising the voltage of node three will not cause any changes before the current from node three switches from a sink to a source. However, if the voltage at node three is decreased then at some

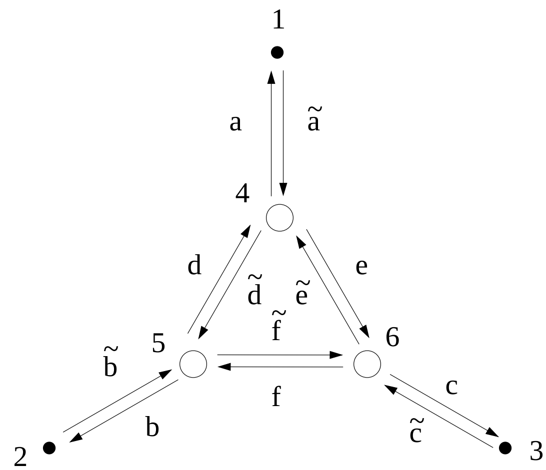

FIGURE 4.1. The delta spike graph with labeled vertices and conductances.

point before node two becomes a source the current on the five six edge will switch direction.

These three cases being accounted for allow us to now detect what boundary voltages causes there to be zero current on the five-six edge. By symmetry a similar technique will allow us to discern when the other two interior-interior edges carry zero current. We now denote the current at boundary node j and the voltage at node q when node p has a voltage of one as i_{pj} and u_{pq} respectively. We gain the following system of equations.

$$\begin{aligned} i_{12} &= b(u_1 2 - u_1 5) \\ i_{13} &= c(u_1 3 - u_1 6) \\ u_{15} &= u_{16} \end{aligned}$$

$$\begin{aligned} i_{23} &= c(u_2 3 - u_2 6) \\ i_{21} &= a(u_2 1 - u_2 4) \\ u_{24} &= u_{26} \end{aligned}$$

$$\begin{aligned} i_{31} &= a(u_3 1 - u_3 4) \\ i_{32} &= b(u_3 2 - u_3 5) \\ u_{34} &= u_{35} \end{aligned}$$

We now combine each set of three equations to yield the following three.

$$u_3 - u_2 = \frac{i_{13}}{c} - \frac{i_{12}}{b}$$
$$u_1 - u_3 = \frac{i_{21}}{a} - \frac{i_{23}}{c}$$
$$u_2 - u_1 = \frac{i_{32}}{b} - \frac{i_{31}}{a}$$

Solving for conductivities yields.

$$\begin{aligned} c &= \tfrac{i_{13}}{u_{13}-u_{12}+i_{12}/b} \\ c &= \tfrac{-i_{23}}{u_{21}-u_{23}-i_{21}/a} \\ a &= \tfrac{-i_{31}}{u_{32}-u_{31}-i_{32}/b} \end{aligned}$$

Notice that when the initial equations for the currents are plugged into these denominators that the result are expressions which are zero only if there is zero current on the boundary spikes, an impossibility since no two edges can have zero current simultaneously unless the boundary potentials are all equal. Combining these equations and solving for a yields.

$$a = \frac{i_{12}i_{23}i_{31} - i_{13}i_{21}i_{32}}{i_{12}i_{23}(u_{32} - u_{31}) - i_{13}i_{32}(u_{21} - u_{23}) - i_{23}i_{32}(u_{13} - u_{12})} \tag{4.1}$$

Solving for $\tilde{a}$ follows in the same manner except that the boundary voltages are multiplied by negative one. Once the boundary spikes conductances are determined the resulting graph can be contracted to a bidirected delta, which is clearly recoverable.

The ring trick can be generalized to a broader class of graphs. Suppose that there are three interior nodes that are all connected to each other by disjoint subgraphs such that the net conductance of each subgraph as viewed by these three interior nodes is different depending on the direction of current. If there is a spike at each one of those interior nodes those spikes can be recovered.

This case is essentially the same as the ring trick for the delta spike. Assign to all boundary nodes other than the three spikes zero-current, (this would be the same as simply unhooking the boundary nodes from power supplies). While at first this seems like a mixed Neumann, Dirichlet problem it can be reduced to a Dirichlet problem because the boundary nodes with zero-current are no different interior nodes, this gives us existence and uniqueness.

In a graph that has only two boundary nodes the only two possible current patterns occur when one node has a higher voltage than the other and when that relationship is the opposite. Thus in our graph where the boundary nodes other than the spikes have zero current the subgraphs between the interior nodes, 4, 5, and 6 will behave exactly like individual strictly bidirectional edges in that they will have one net conductance in one direction and different in the opposite direction, only carrying zero current if there is no voltage difference.

This technique can be used recover graphs such as the spiked square. There also exists a very similar trick that can be used to recover directed graphs that aren't

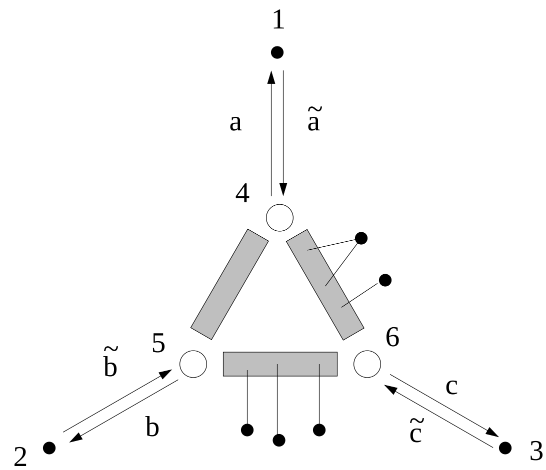

FIGURE 4.2. The generalized delta spike graph where the edges of the delta are subgraphs.

required to be bidirected. So far discontinuity tricks more complicated than the delta spike trick have been shown to be equivalent to the delta spike trick.

5. Future Work

There are a number of natural questions raised by this topic and paper.

- A proof of conjecture 2.2 would be nice. My thoughts on that subject are in that section.
- More complicated analysis of directed graphs will find some more complicated method of using discontinuities to determine internal properties of the network. The problem has so far been that the information gained from the discontinuities has only been useful in the cases where the location of the zero current can be determined from the topology of the graph. Based on the intuition gained from the undirected case I suspect that determining the location of all of the discontinuities will reveal something even if each one individually does not. This line of thought though may require a stronger a statement of the inverse problem, where perhaps the response on the whole n-2 sphere is given.
- A very closely related topic is the study of directed graphs in general. In some cases these graphs can be recovered using a modified version of the recovery algorithm laid out in this paper. The modification is that at instead of associating a single undirected graph to each point in a region 2^E graphs must now be associated, one for all possible combinations of an edge carrying zero current. The solutions of each region are kept until all regions are solved for. Since there exists a solution at least on set of conductances from each region will be consistent. In general though that only works rarely and is rather ugly. It also seems that the conditions on edge orientations that allow for a graph to be recovered are complex. Being able to detect discontinuities is again a useful tool and may be more valuable in the general directed than in the bidirected case. In fact the

directed delta, when it is recoverable, can only be recovered by using a discontinuity trick. There is a lot to be figured out about directed graphs in general.

- The idea of constructing a point response matrix could be used to recover nonlinear networks as well. Such a method would work for recovering graphs where conductances varied with the square of the voltage difference or other nonlinear situations.

References

[1] Orion Bawden., Paper pending

[2] Curtis, Edward B. & Morrow, James A., *Inverse Problems for Electrical Networks*, Series on Applied Mathematics - Vol. 13, World Scientific, New Jersey, 2000.