Transfer of status qualifying dissertation:
a compositional approach to control theory

Brendan Fong

Department of Computer Science
University of Oxford
United Kingdom OX1 3QD

brendan.fong@cs.ox.ac.uk

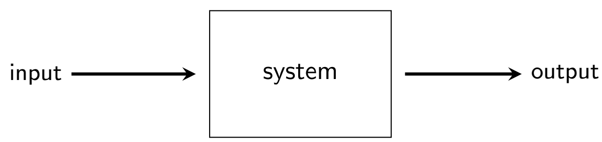
December 16, 2013

Abstract

This paper proposes a programme of research with the aim of understanding network-based systems through the application of monoidal category theory, and outlines some in-progress work in this direction.

1 Introduction

Control theory begins with the following picture:



In this picture we have an object under study, referred to as a *system*, which when fed certain inputs produces certain outputs. These outputs need not be uniquely determined by the inputs; in general the relationship may be stochastic or non-deterministic, or depend also on internal states of the system. Although we take it as given that the system is an *open* system—so it interacts with its environment, and so we can observe its inputs and outputs—we assume no access to the details of these internal states or the inner workings of the system, and this can add considerable complexity to our models. The end goal of control theory is then to control the system: to understand how to influence its behaviour in order to achieve some desired goal. Two key questions arise:

- **Analysis**: Given a system, what is the relationship it induces between input and output?
- **Synthesis**: Given a target relationship between input and output, how do we build a system that produces this relationship?

The first, the question of system analysis, provides the basic understanding required to find the inputs that lead to the desired outputs. The second question, the question of synthesis, is the central question of feedback control theory, which aims to design controllers to regulate other systems. We give brief overview of the field, illustrated with some questions of these kinds.

Control has a long history. Indeed, many systems found in the biology and chemistry of living organisms have interesting interpretations from a control theoretic viewpoint, such as those that are responsible for body temperature regulation or bipedal balance [39]. Human understanding of control developed alongside engineering, and so dates back at least as far as antiquity, with for example the ancient Romans devising elaborate systems to maintain desired water levels in aqueducts [38]. The origin of formal mathematical control theory, however, is more recent, and in general taken to be James Clerk Maxwell's seminal analysis of centrifugal governors, presented to the Royal Society of London in 1868 [12].

A good example of a basic, but characteristic, system studied by control theory, centrifugal governors were developed to address the problem of speed regulation for steam engines. A centrifugal governor consists of masses attached via lever arms to a vertical spindle, as depicted in Figure 1. This spindle is then connected to the output wheel of an engine, and the height of the masses is inversely related to the aperture of the steam valve into the engine. Higher speeds therefore cause the spindle to rotate faster, which in turn causes the masses to move upward and decrease the flow of steam into the engine. This moderates or prevents acceleration, and hence overspeeding of the engine. Maxwell's paper classifies the possible trajectories over time of the speed of a governed steam engine.

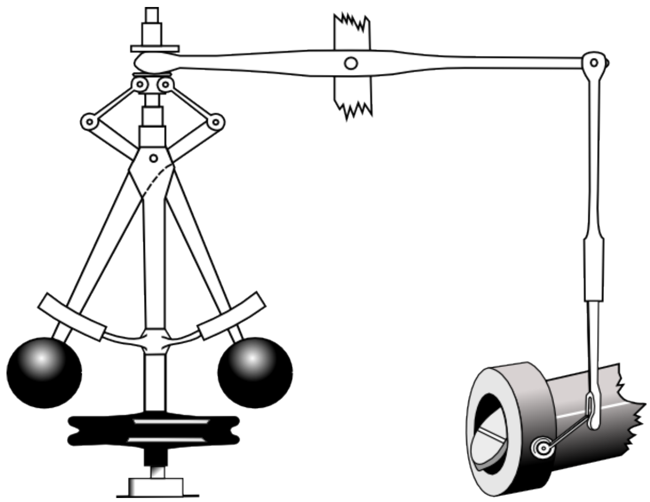


Figure 1: A centrifugal governor.

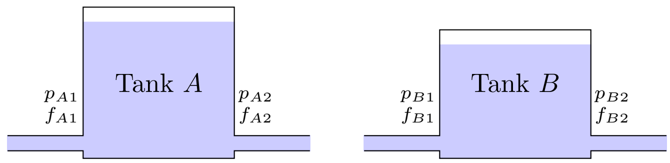
With respect to the framework laid out above, there are three systems under investigation here. The first is the steam engine, with input steam flow and output engine speed. The second is the centrifugal governor, with input engine speed and output steam flow. The third is the combined governor-steam engine, which forms the main system under investigation, with again input steam flow and output engine speed.

The techniques used and developed from Maxwell's paper, in particular by Rayleigh and Heaviside, are in general known as *transfer function* techniques. A transfer function is a linear map from the set of inputs to the set of outputs of a linear time-invariant system with zero initial conditions. Transfer functions are commonly used in the analysis of single-input single-output linear systems, but become unwieldy or inapplicable for more general systems.

In the 1960s, led by Wiener and Kalman, so-called *state space* techniques were developed to address multiple-input multiple-output, time-varying, nonlinear systems [14]. These methods are characterised by defining a system as a collection of input, output, and internal state variables, with the state changing as a function of the input and state variables over time, and the output a function of the input and state. These functions are often implicitly specified by differential equations.

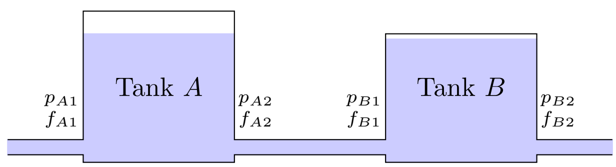
In general, however, classical control theory remains grounded in a paradigm that defines and analyses systems in terms of inputs and outputs or, from another perspective, causes and effects. In recent years Willems, among others, has argued that this input-output perspective is limiting, as in the case of many systems studied through control theory there is no clear distinction between input and output variables [47]. For example, given a circuit component for which the relevant variables are the voltages and currents, different contexts may call for the voltage to be viewed as the input and current output, or vice versa. It is useful to have a single framework capable of discussing the behaviour of the component without making this choice.

Moreover, in the drive to understand larger and more complex systems, increasing emphasis has been put on understanding the way systems can be broken down into composite subsystems, and conversely how systems link together to form larger systems [48, 20]. Indeed, interconnection or composition of systems has always played a central role in systems engineering, as in the combination of the centrifugal governor and steam engine of Maxwell's research, and the difficulty of discussing how systems compose within an input-output framework lends support to Willems' call for a more nuanced definition of control system. To illustrate these difficulties, consider the simple example of two water tanks each with two access pipes:



For each tank, the relevant variables are the pressure p and the flow f. A typical control theoretic analysis might view the water pressure as inducing flow through the tank, and hence take pressure as the input variable, flow as the output variable, and describe each tank as a transfer function $H_\bullet : P_\bullet \to F_\bullet$ between the sets $P_\bullet$ and $F_\bullet$ these vary over.

Ideally then, the composite system below, connecting pipe 2 of Tank A to pipe 1 of Tank B, would be described by the composite of the transfer functions H_A and H_B of these tanks.



This is rarely the case, and indeed makes little sense: the output of transfer function H_A describes the flow through Tank A, while the domain of the transfer function H_B describes pressures through Tank B, so taking the output of H_A as the input of H_B runs into data type issues. Instead, here the relationship between the transfer functions H_A and H_B, and the transfer function H_{AB} of the composite system can be understood through the fact that connection requires the pressure at pipe 1 of Tank A must be equal to the pressure at pipe 2 of Tank B, and that the flow out of the former pipe must equal the flow into the latter—that is, by imposing the relations

$$p_{A2} = p_{B1}$$
$$f_{A2} = -f_{B1}$$

on the variables. Indeed in many contexts, hydraulics and electronics among them, connections between systems are characterised not by the 'output' of one system forming the 'input' of the next, but by *variable sharing* between the systems. Such relations are often difficult to describe using the language of transfer functions and state space methods.

This dissertation proposes to contribute to a programme of research that aims to place composition of systems as central to the study of control, and in doing so address the limitations of the input-output paradigm that these two dominant approaches to control theory fall within.

The following section, Section 2, gives a brief overview of the context this programme arises from. Section 3 then runs through an example of the intended approach, and Section 4 proposes a few intended directions for further study.

2 Literature review

2.1 Behaviours

We take as our starting point a perspective of Willems', which models control systems as collections of possible parameter readings for the variables of the system [29]. More formally, we model systems using triples (T, W, B) comprising a time axis $T \subseteq \mathbb{R}$, a signal space W in which the variables take their values, and a **behaviour** $B \subseteq W^T$. The behaviour is thought of as describing the trajectories of parameter readings over time that are declared possible by the system.

In cases where transfer function or state space techniques may be applied, the behaviour is often the set of ordered pairs defining the transfer function, or the set of solutions to a system of differential equations. Note that this perspective also allows for discussion of systems that are not usually thought of in such ways, such as a tank of gases governed by the ideal gas law, or a collection of planets governed by Kepler's law. Indeed, in general physical laws only prescribe a set of allowed trajectories: they make no comment on cause and effect. Thus a framework that does not assume these ideas may be of use.

2.2 Networks, effort, and flow

Phrased in these terms, the two central questions of control theory posed above can be viewed as investigating the properties of two opposing maps:

$$\textbf{Systems} \underset{\text{synthesis}}{\overset{\text{analysis}}{\rightleftarrows}} \textbf{Behaviours}$$

These questions only become well-defined within the context of a specified class of systems, such as hydraulic systems or mechanical structures or electrical circuits. In the work of this proposed dissertation, we shall restrict attention to contexts in which notions of effort and flow are present in graph or network-like structures, indeed such as in hydraulic systems or electrical circuits, and certain mechanical constructions. First discussed in detail by Olsen [28], mathematically precise analogies exist between, for example, systems of the types given by the following table, and their associated quantities.

	displacement q	flow q'	momentum p	effort p'
Translation	position	velocity	momentum	force
Rotation	angle	angular velocity	angular momentum	torque
Electronics	charge	current	flux linkage	voltage
Hydraulics	volume	flow	pressure momentum	pressure
Thermodynamics	entropy	entropy flow	temperature momentum	temperature
Chemistry	moles	molar flow	chemical momentum	chemical potential

Further work, pioneered by in particular Odum, links these network-type systems to biology, ecology, and economics, and we anticipate results of this research programme making a contribution across all these fields [27].

2.3 A diagrammatic approach

A result of these analogies is that many of the diagrammatic formalisms designed to reason about these systems, such as electric and hydraulic circuit diagrams (Figure 2), signal flow graphs (Figure 3), and energy systems language (Figure 4), themselves have many visual and deeper similarities. Bond graphs, originated by Paynter in 1961 [30] and further developed and applied by Karnopp,

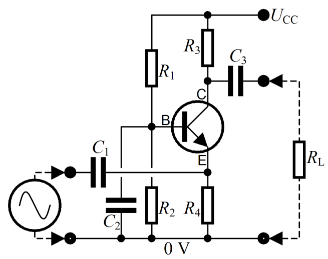


Figure 2: An electrical circuit representing an amplifier.

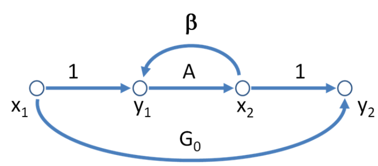


Figure 3: A signal flow graph for the asymptotic gain model for a negative feedback amplifier.

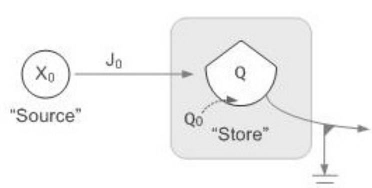


Figure 4: An energy systems diagram depicting an energy storage system.

Rosenberg, and Thoma, among others [9, 42], go some way to unifying these diagrammatic languages into a common, domain-independent framework.

In these formalisms each diagram represents a system. The diagrammatic language itself comprises a collection of atomic symbols, a syntax for connecting them to result in 'well-formed' diagrams, and hence a well-defined representation of a system, and a rule for understanding the semantics of the diagrams, for example associating a transfer function or differential equation to each diagram.

The independent emergence and widespread use of diagrammatic methods across these fields suggests that diagrams are an effective way to understand this type of system, and we take this as a guiding principle. In particular, a strength of a diagrammatic approach is that the syntax for these diagrams gives rise to an understanding of how to connect diagrams. While this is so, however, within the context of transfer functions or state space methods it is often difficult to understand how the semantics of the combined system relates to the semantics of the constituent systems. We suggest that using behaviours to understand systems, and so the semantics of these diagrams, allows for a better understanding of how the semantic interpretations of the diagrams compose.

2.4 Dagger compact categories

Past work on formalising these diagrammatic languages has often made use of the language of graph theory, treating diagrams as graphs with leaves [47]. In line with our emphasis on the compositional aspects of control theory, we propose the use of category theory. The two-dimensional nature of these graphical formalisms suggest monoidal structure is useful.

Monoidal categories were first defined by Mac Lane in 1963 [24]. Arising from the work of Penrose [31] and Joyal and Street [18, 19], it became clear that morphisms in these categories had a precise graphical representation through string diagrams, and subsequently that monoidal categories can be used to provide a foundation for Feynman diagrams and other similar graphical tools.

We shall be in particular concerned with dagger compact categories—symmetric monoidal categories further equipped with a dagger functor and duals for objects. Recent work including that by Abramsky and Coecke [1], Baez [3], and Selinger [35] has shown this is an important structure in quantum mechanics. The dagger compact structure here reflects the interchangability between the domain and codomain of a diagram and our acausal view of systems; we give an example in the following section.

We note also related work by Spivak on the operad of wiring diagrams [40].

2.5 Symplectic methods

While the diagrams themselves give a straightforward dagger compact category of systems, we aim also to describe the behaviours of a given class of systems as a dagger compact category. As the behaviours of a system can often be described as those that minimise a certain energy, techniques from symplectic geometry are useful here. In particular, drawing from the work of Hörmander and Sniatycki and Tulczejew, we shall see that it is possible to describe behaviours as Lagrangian submanifolds of symplectic manifolds, and that these form a category [46].

The role of in particular this category and more generally of symplectic geometry in control theory has been explored by Benenti [8], while Cimasoni and Turaev have done work on the relationship between categories of diagrams and categories of Lagrangian submanifolds [13]

3 Pilot work

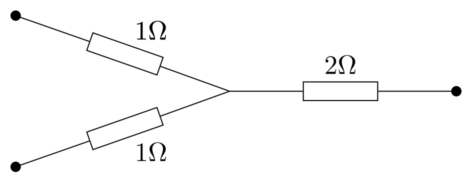
To demonstrate the viability of this project, in this section we briefly explore, omitting proofs, how these ideas develop in the context of electrical circuits consisting only of linear resistors. We begin with the construction of a category whose morphisms are these networks of linear resistors, before constructing a category classifying the possible behaviours of these systems, and then discussing the 'analysis' functor mapping systems to behaviours.

Much here is owed to unpublished work and discussions with John Baez.

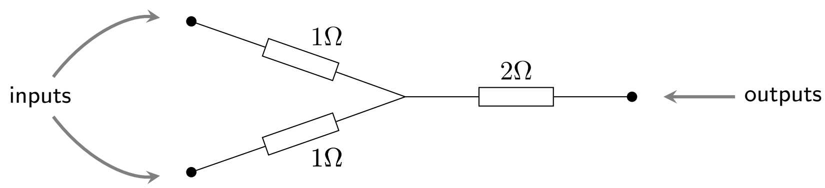
3.1 Circuits of linear resistors

The concept of an abstract electrical circuit made of linear resistors is well-known in electrical engineering, but we shall to formalise it with more precision than usual. The basic idea is that a circuit of linear resistors is a graph whose edges are labelled by positive real numbers called 'resistances', and whose sets of vertices is equipped with two subsets: the 'inputs' and 'outputs'. This unfolds as follows.

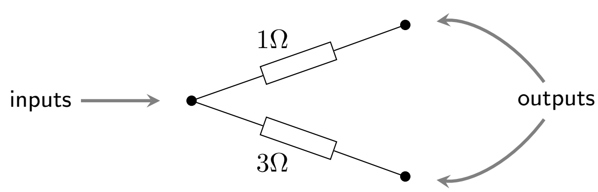
A circuit of resistors looks like this:



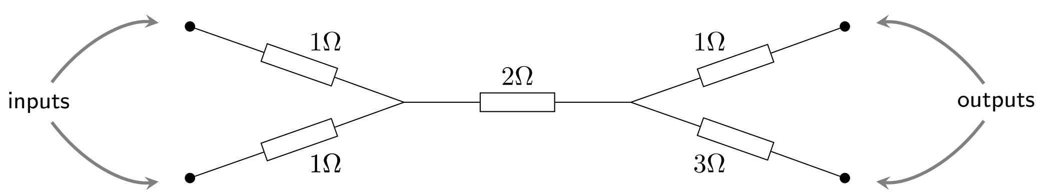
We can consider this a labelled graph, with each resistor an edge of the graph, its resistance its label, and the vertices of the graph the points at which resistors are connected. Composition of circuits involves creating connections between distinct circuits. To do this we first mark points at which connections can be made by denoting some vertices as input and output terminals:



Then, given a second circuit, we may choose a relation between the output set of the first and the input set of this second circuit, such as the simple relation of the single output vertex of the circuit above with the single input vertex of the circuit below.



We compose the two circuits by identifying output and input vertices according to this relation, giving in this case the composite circuit:



We will formalise these notions of 'marking terminals' and composition of circuits using cospans and push forwards. Treating these cospans as morphisms between finite sets, we shall show that this defines a dagger compact category.

3.1.1 L-labelled graphs

Given a set of labels L, we shall define a category L-Graph of L-labelled directed multigraphs. In this paper we will refer to directed multigraphs simply as graphs. So, define a **graph** to be a pair of functions $s, t : E \to N$ where E and N are finite sets. An L-**graph** is a graph further equipped with a function $r : E \to L$.

$$\begin{tikzcd} L & E \arrow[l, "r"'] \arrow[r, shift left, "s"] \arrow[r, shift right, "t"'] & N \end{tikzcd}$$

We call elements of E **edges** and elements of N **vertices** or **nodes**. We say that the edge $e \in E$ has **source** $s(e)$ and **target** $t(e)$, and also say that e is an edge **from** $s(e)$ **to** $t(e)$.

Given L-graphs $\Gamma = (E, N, s, t, r)$ and $\Gamma' = (E', N', s', t', r')$, a morphism of L-graphs $\Gamma \to \Gamma'$ is a pair of functions ϵ, v such that the following diagrams commute:

$$\begin{tikzcd} & E \arrow[dl, "r"'] \arrow[dd, "\epsilon"] \\ L & \\ & E' \arrow[ul, "r'"] \end{tikzcd} \qquad \begin{tikzcd} E \arrow[r, "s"] \arrow[d, "\epsilon"'] & N \arrow[d, "v"] \\ E' \arrow[r, "s'"'] & N' \end{tikzcd} \qquad \begin{tikzcd} E \arrow[r, "t"] \arrow[d, "\epsilon"'] & N \arrow[d, "v"] \\ E' \arrow[r, "t'"'] & N'. \end{tikzcd}$$

These L-graphs and their morphisms form the category ***L*-Graph**. Using results about colimits in the category of sets, it is straightforward to check that this category has finite colimits.

We will later make use of the notion of connectedness in graphs, and so we give a definition here. Given two vertices $v, w \in N$ of a graph, a **path from v to w** is a finite sequence of vertices $v = v_0, v_1, \dots, v_n = w$ and edges $e_1, \dots, e_n$ such that for each $1 \leq i \leq n$, either e_i is an edge from v_i to v_{i+1}, or an edge from v_{i+1} to v_i. A subset S of the vertices of a graph is **connected** if for each pair of vertices in S, there is a path from one to the other. A **connected component** of a graph is a maximal connected subset of its vertices.

3.1.2 Cospans

In this paper we are particularly interested in the case when our set L of labels is the set $(0, \infty)$ of positive real numbers. Here we view each edge as a 'wire' between the source and target vertices, and consider each label as assigning a 'resistance' to the wire, so that the objects of $(0, \infty)$-Graph may be considered as networks of resistors. Cospans provide the machinery for combining these networks in series.

Recall that a **cospan** from X to Y in a category $\mathcal{C}$ is an object C in $\mathcal{C}$ with a pair of morphisms $f : X \to C$, $g : Y \to C$.

$$\begin{tikzcd} & C & \\ X \arrow[ur, "f"] & & Y. \arrow[ul, "g"'] \end{tikzcd}$$

We shall refer to X and Y as the **feet**, and C as the **apex** of the cospan. When such pushforwards exist, cospans may be composed using the pushforward from the common foot: given cospans $(C, f : X \to C, g : Y \to C)$ for X to Y and $(C', f' : Y \to C', g' : Z \to C')$ from Y to Z, their composite cospan is $(P, i \circ f : X \to P, i' \circ g' : Z \to P)$, where $(P, i : C \to P, i' : C' \to P)$ is the pushforward of the square:

$$\begin{tikzcd} & & P & & \\ & C \arrow[ur, "i"] & & C' \arrow[ul, "i'"'] & \\ X \arrow[ur, "f"] & & Y \arrow[ul, "g"'] \arrow[ur, "f'"] & & Z. \arrow[ul, "g'"'] \end{tikzcd}$$

A **map of cospans** is a morphism $h : C \to C'$ in $\mathcal{C}$ between the apices of two cospans $(C, f : X \to C, g : Y \to C)$ and $(C', f' : X \to C', g' : Y \to C')$ with the same feet, such that

$$\begin{tikzcd} & C \arrow[dd, "h"] & \\ X \arrow[ur, "f"] \arrow[dr, "f'"'] & & Y \arrow[ul, "g"'] \arrow[dl, "g'"] \\ & C' & \end{tikzcd}$$

commutes.

We will later discuss **spans** in a category $\mathcal{C}$. The definitions of these, their composition, and their maps are dual to the corresponding definitions above.

3.1.3 A category of circuits

Following the work of Stay, L-labelled graphs, cospans in L-Graph, and maps of cospans form a compact closed bicategory **Cospan(L-Graph)** [41]. Note that this requires choosing pushforwards for each pushforward square in L-Graph. We shall predominantly concern ourselves with the compact closed subbicategory ResCirc of Cospan($(0,\infty)$-Graph), which we define as follows.

Definition 3.1. *The bicategory* **ResCirc** *is the full subbicategory of* Cospan($(0,\infty)-$Graph) *with objects those* $(0,\infty)$*-graphs with no edges. We call the 1-morphisms of* ResCirc **circuits**.

It can be shown that every object in Cospan($(0,\infty)$-Graph) is self-dual, and this implies that ResCirc is again a compact closed bicategory [41].

Since $(0,\infty)$-graphs with no edges are uniquely determined by their vertex sets, an object of ResCirc may simply be thought of as a finite set, while a 1-morphism between finite sets X and Y may be thought of as comprising a $(0,\infty)$-graph $\Gamma = (E, N, s, t, r)$ with functions $i : X \to N$ and $o : Y \to N$. We call the sets X and Y the **input connections** and **output connections** of the circuit respectively, and call subsets $N_- = i(X)$ and $N_+ = o(Y)$ the **input terminals** and **outputs terminals** of the circuit. These are both thought of as points to which we might attach connections or terminals of other circuits. Note that N_- and N_+ need not be disjoint. Often the difference between inputs and outputs will not matter, so we also define $\partial N = N_- \cup N_+$, and collectively call elements of this set **terminals**.

Abusing notation slightly, we also write ResCirc for the 1-categorical decategorification of this bicategory, with objects finite sets, and morphisms isomorphism classes of cospans in $(0,\infty)$-Graph with feet such objects. Note that the notion of isomorphism for cospan of $(0,\infty)$-Graph is quite restrictive, amounting to no more than a renaming of the nodes and edges, so we may be lazy with our distinction between the notions of morphism of this category and of circuits.

This category is compact closed, with dual objects and their units and counits given by the compactness of the bicategory. The compactness of our category of circuits of linear resistors captures the interchangeability between individual input and output terminals of circuits—that is, the fact that we can choose any input terminal to our circuit and consider it instead as an output terminal, and vice versa.

Exploring the structure more explicitly, the monoidal product for ResCirc is simply the disjoint union of finite sets on objects, and the disjoint union of graphs for morphisms. The monoidal unit is the empty set. Observe that given a morphism $f : X \to Y$ in some category $\mathcal{C}$, we may construct the cospan

$$\begin{tikzcd} & Y & \\ X \arrow[ur, "f"] & & Y \arrow[ul, "\mathrm{id}_Y"'] \end{tikzcd}$$

from X to Y, with this cospan an isomorphism if and only if f is, and with inverse

$$\begin{tikzcd} & Y & \\ Y \arrow[ur, "\mathrm{id}_Y"] & & X \arrow[ul, "f"'] \end{tikzcd}$$

in this case. The structural morphisms for the symmetric monoidal structure may be specified by set functions in this way, with for example the associator the cospan induced by the canonical functions $(X \sqcup Y) \sqcup Z \to X \sqcup (Y \sqcup Z)$, and the braiding the cospan induced by the canonical functions $X \sqcup Y \to Y \sqcup X$.

The existence of duals may be seen as follows. Let X be a finite set, and let $a : X \sqcup X \to X$ be the function given by taking the coproduct of two identity functions $X \to X$, and let $! : \varnothing \to X$ be the unique such map. We may take X as its own dual object, with unit and counit given by

$$\begin{tikzcd} & X & \\ \varnothing \arrow[ur, "!"] & & X \sqcup X \arrow[ul, "a"'] \end{tikzcd} \qquad \text{and} \qquad \begin{tikzcd} & X & \\ X \sqcup X \arrow[ur, "a"] & & \varnothing. \arrow[ul, "!"'] \end{tikzcd}$$

It is readily verified that these cospans do indeed form a unit and counit for the self-dual object X.

3.1.4 The dagger structure

A dagger functor expresses the idea that the direction of morphisms can be reversed: through a dagger functor each morphism specifies a map from its codomain to its domain, in addition to the map it is from its domain to its codomain. This is true of electrical circuits: if we like we may treat the set of inputs as the set of outputs instead, and the set of outputs as the set of inputs.

Proposition 3.2. *The category* ResCirc *is a dagger compact category.*

3.2 Equivalences of circuits, and Dirichlet forms

So far we have modelled our circuits of linear resistors according to their physical form. But another, often more relevant, way to understand a circuit is by its function. To an electric circuit we associate two quantities to each edge: voltage and current. However, circuits are subject to governing laws that imply these quantities must have certain relationships; we are not free to choose voltages and currents as we like. From the perspective of control theory we are particularly interested in the values these quantities take at the terminals, and how altering one value will affect the other values. We shall see that the potentials at the terminals of the circuit determine the flow of current in or out of these terminals. We call two circuits equivalent when they determine the same relationship. Our task in this section is to explore when two circuits are equivalent. We begin by discussing how to find the function of a circuit from its form, advocating in particular the perspective of the principle of minimum power.

In the following let X and Y be finite sets, and $\Gamma = (E, N, s, t, r)$, together with maps $i : X \to N$ and $o : Y \to N$, be a circuit. In particular E will always stand for the edge set of the apex of a circuit, and N for the set of nodes. For a finite set S, we shall write $\mathbb{R}^S$ for the vector space of functions $\psi : S \to \mathbb{R}$. Elementary circuit theory is most frequently presented in terms of voltages and currents on the edges or 'wires' of the circuit. These are functions $V \in \mathbb{R}^E$ and $I \in \mathbb{R}^E$ respectively. We use I for 'intensity of current', following Ampère. This material is all well known, and the current presentation is a rewriting of a draft by Baez [4].

3.2.1 Ohm's law and Kirchhoff's laws

We say that **Ohm's law** holds if for all edges $e \in E$ the voltage and current functions of a circuit obey:

$$V(e) = r(e)I(e).$$

We say **Kirchhoff's voltage law** holds if there exists $\phi \in \mathbb{R}^N$ such that

$$V(e) = \phi(t(e)) - \phi(s(e)).$$

We call the function ϕ a **potential**, and think of it as assigning an electrical potential to each node in the circuit. The voltage then arises as the differences in potentials between adjacent nodes. If Kirchhoff's voltage law holds for some voltage V, the potential ϕ is unique only in the trivial case of the empty circuit: when the set of nodes N is empty. Indeed, two potentials define the same voltage function if and only if their difference is constant on each connected component of the graph Γ.

We say **Kirchhoff's current law** holds if for all nonterminal nodes $n \in N \setminus \partial N$ we have

$$\sum_{s(e)=n} I(e) = \sum_{t(e)=n} I(e).$$

This is an expression of conservation of charge within the circuit; it says that the total current flowing in or out of any nonterminal node is zero. Even when Kirchhoff's current law is obeyed, terminals need not be sites of zero net current; we call the function $\iota \in \mathbb{R}^{\partial N}$ that takes a terminal to the difference between the outward and inward flowing currents,

$$\begin{aligned} \iota : \partial N &\longrightarrow \mathbb{R} \\ n &\longmapsto \sum_{t(e)=n} I(e) - \sum_{s(e)=n} I(e), \end{aligned}$$

the **boundary current** for I.

3.2.2 The principle of minimum power

A **boundary potential** is also a function in $\mathbb{R}^{\partial N}$, but instead thought of as specifying electric potentials on the terminals of a circuit. As we think of our circuits as open circuits, with the terminals points of interaction with the external world, we shall think of these potentials as variables that are free for us to choose. Using the above three principles—Ohm's law, Kirchhoff's voltage law, and Kirchhoff's current law—it is possible to show that choosing a boundary potential determines unique voltage and current functions on that circuit.

The principle of minimum power gives some insight into how this occurs, by describing a way potentials on the terminals might determine potentials at all nodes. From this, Kirchhoff's voltage law then gives rise to a voltage function on the edges, and Ohm's law gives us a current function too. In fact, a potential satisfies the principle of minimum power for a given boundary potential if and only if this current obeys Kirchhoff's current law.

A circuit with current I and voltage V dissipates energy at a rate proportional to the real number

$$P = \sum_{e \in E} I(e)V(e).$$

Ohm's law allows us to rewrite I as V/r, while Kirchhoff's voltage law gives us a potential ϕ such that $V(e)$ can be written as $\phi(t(e)) - \phi(s(e))$, so for a circuit obeying these two laws the power can also be expressed in terms of this potential. We thus arrive at a functional mapping potentials ϕ to the power dissipated by the circuit when Ohm's law and Kirchhoff's voltage law are obeyed for ϕ.

Definition 3.3. *The* **extended power functional** $P : \mathbb{R}^N \to \mathbb{R}$ *of a circuit is the map*

$$P(\varphi) = \sum_{e \in E} \frac{1}{r(e)} \big(\varphi(t(e)) - \varphi(s(e))\big)^2.$$

We call this the *extended* power functional as we shall see that it is defined on potentials that are not compatible with the three governing laws of electric circuits. We shall later restrict the domain of this functional so that it is defined precisely on those potentials that *are* compatible with the governing laws. Note that this functional does not depend on the directions chosen for the edges of the circuit.

This expression lets us formulate the 'principle of minimum power', which gives us information about the potential ϕ given its restriction to the boundary of Γ. Call a potential $\phi \in \mathbb{R}^N$ an **extension** of a boundary potential $\psi \in \mathbb{R}^{\partial N}$ if ϕ is equal to ψ when restricted to $\mathbb{R}^{\partial N}$—that is, if $\phi|_{\mathbb{R}^{\partial N}} = \psi$.

Definition 3.4. *We say a potential* $\phi \in \mathbb{R}^N$ **obeys the principle of minimum power** *for a boundary potential* $\psi \in \mathbb{R}^{\partial N}$ *if* ϕ *minimizes the extended power functional* P *subject to the constraint that* ϕ *is an extension of* ψ.

In the presence of Ohm's law and Kirchhoff's voltage law, it can be shown that the principle of minimum power is equivalent to Kirchhoff's current law.

Proposition 3.5. *Let* ϕ *be a potential extending some boundary potential* ψ. *Then* ϕ *obeys the principle of minimum power for* ψ *if and only if the induced current* $I(e) = \frac{1}{r(e)}(\phi(t(e)) - \phi(s(e)))$ *obeys Kirchhoff's current law.*

3.2.3 A Dirichlet problem

We remind ourselves that we are in the midst of understanding circuits as objects that define relationships between boundary potentials and boundary currents. This relationship is defined by the stipulation that voltage–current pairs on a circuit must obey Ohm's law and Kirchhoff's laws—or equivalently, Ohm's law, Kirchhoff's voltage law, and the principle of minimum power. In this subsection we show these conditions imply that for each boundary potential ψ on the circuit there exists a potential ϕ on the circuit extending ψ, unique up to what may be interpreted as a choice of reference potential on each connected component of the circuit. From this potential ϕ we can then compute the unique voltage, current, and boundary current functions compatible with the given boundary potential.

Fix again a circuit with extended power functional $P : \mathbb{R}^N \to \mathbb{R}$. Let $\nabla : \mathbb{R}^N \to \mathbb{R}^N$ be the operator that maps a potential $\phi \in \mathbb{R}^N$ to the function $N \to \mathbb{R}$ given by

$$n \longmapsto \left.\frac{\partial P}{\partial \varphi(n)}\right|_{\varphi = \phi}.$$

As we have seen, this function takes potentials to twice the pointwise currents that they induce. We have also seen that a potential ϕ is compatible with the governing laws of circuits if and only if

$$\big(\nabla\phi\big)\big|_{\mathbb{R}^{\partial N}} = 0 \tag{1}$$

The operator ∇ acts as a discrete analogue of the Laplacian for the graph Γ, so we call this operator the **Laplacian** of Γ, and say that the equation (1) is a version of Laplace's equation. We then say that the problem of finding an extension ϕ of some fixed boundary potential ψ that solves this Laplace's equation—or, equivalently, the problem of finding a ϕ that obeys the principle of minimum power for ψ—is a discrete version of the **Dirichlet problem**.

As we shall see, this version of the Dirichlet problem always has a solution. However, the solution is not necessarily unique. If we take a solution ϕ and some $\alpha \in \mathbb{R}^N$ that is constant on each connected component and vanishes on the boundary of Γ, it is clear that $\phi + \alpha$ is still an extension of ψ and that $\frac{\partial P}{\partial\varphi(n)}|_{\varphi=\phi} = \frac{\partial P}{\partial\varphi(n)}|_{\varphi=\phi+\alpha}$, so $\phi + \alpha$ is another solution. We say that a connected component of a circuit **touches the boundary** if it contains a vertex in ∂N. Note that such an α must vanish on all connected components touching the boundary.

With these preliminaries in hand, we can solve the Dirichlet problem:

Proposition 3.6. *For any boundary potential $\psi \in \mathbb{R}^{\partial N}$ there exists a potential ϕ obeying the principle of minimum power for ψ. If we also demand that ϕ vanish on every connected component of Γ not touching the boundary, then ϕ is unique.*

Proof. For existence, observe that the power is a nonnegative quadratic form, the extensions of ψ form an affine subspace of $\mathbb{R}^N$, and a nonnegative quadratic form restricted to an affine subspace of a real vector space must reach a minimum somewhere on this subspace.

For uniqueness, suppose that both ϕ and ϕ' obey the principle of minimum power for ψ. Let

$$\alpha = \phi' - \phi.$$

Then

$$\alpha|_{\partial N} = \phi'|_{\partial N} - \phi|_{\partial N} = \psi - \psi = 0,$$

so $\phi + \lambda\alpha$ is an extension of ψ for all $\lambda \in \mathbb{R}$. This implies that

$$f(\lambda) := P(\phi + \lambda\alpha)$$

is a smooth function attaining its minimum value at both $t = 0$ and $t = 1$. In particular, this implies that $f'(0) = 0$. But this means that when writing f as a quadratic, the coefficient of λ must be 0, so we can write

$$\begin{aligned}
f(\lambda) &= \sum_{e\in E} \frac{1}{r(e)}\big((\phi + \lambda\alpha)(t(e)) - (\phi + \lambda\alpha)(s(e))\big)^2 \\
&= \sum_{e\in E} \frac{1}{r(e)}\Big(\big(\phi(t(e)) - \phi(s(e))\big) + \lambda\big(\alpha(t(e)) - \alpha(s(e))\big)\Big)^2 \\
&= \sum_{e\in E} \frac{1}{r(e)}\big(\phi(t(e)) - \phi(s(e))\big)^2 + \lambda\text{-term} + \lambda^2 \sum_{e\in E} \frac{1}{r(e)}\big(\alpha(t(e)) - \alpha(s(e))\big)^2 \\
&= \sum_{e\in E} \frac{1}{r(e)}\big(\phi(t(e)) - \phi(s(e))\big)^2 + \lambda^2 \sum_{e\in E} \frac{1}{r(e)}\big(\alpha(t(e)) - \alpha(s(e))\big)^2.
\end{aligned}$$

Then

$$f(1) - f(0) = \sum_{e\in E} \frac{1}{r(e)}\big(\alpha(t(e)) - \alpha(s(e))\big)^2 = 0,$$

so $\alpha(t(e)) = \alpha(s(e))$ for every edge $e \in E$. This implies that α is constant on each connected component of the graph Γ of our circuit.

Note that as $\alpha|_{\partial N} = 0$, α vanishes on every connected component of Γ touching the boundary. Thus, if we also require that ϕ and ϕ' vanish on every connected component of Γ not touching the boundary, then $\alpha = \phi' - \phi$ vanishes on all connected components of Γ, and hence is identically zero. Thus $\phi' = \phi$, and this extra condition ensures a unique solution to the Dirichlet problem. □

Note also from the proof of the above proposition that:

Proposition 3.7. *Suppose $\psi \in \mathbb{R}^{\partial N}$ and ϕ is a potential obeying the principle of minimum power for ψ. Then ϕ' obeys the principle of minimum power for ψ if and only if the difference $\phi' - \phi$ is constant on every connected component of Γ and vanishes on every connected component touching the boundary of Γ.*

3.2.4 Behaviours

We have seen that boundary potentials determine, essentially uniquely, the value of all the electric properties across the entire circuit. But from the perspective of control theory, this internal structure is irrelevant: we can only access the circuit at its terminals, and hence only need concern ourselves with the relationship between boundary potentials and boundary currents. In this section we streamline our investigations above to state the precise way in which boundary currents depend on boundary potentials. In particular, we shall see that the relationship is completely captured by the functional taking boundary potentials to the minimum power used by any extension of that boundary potential. Furthermore, each such power functional determines a different boundary potential–boundary current relationship, and so we can conclude that two circuits are equivalent if and only if they have the same power function. A 'behaviour' is an equivalence class of circuits, where two are considered equivalent when the boundary current is the same function of the boundary potential.

First let us check that the boundary current is a function of the boundary potential. For this we introduce an important quadratic form on the space of boundary potentials:

Definition 3.8. *The* **(scaled) power functional** $Q : \mathbb{R}^{\partial N} \to \mathbb{R}$ *of a circuit with extended power functional P is given by*

$$Q(\psi) = \frac{1}{2} \min_{\phi|_{\mathbb{R}^{\partial N}} = \psi} P(\phi).$$

Proposition 3.6 shows the minimum above exists, so the power functional is well-defined. Up to a factor of $\frac{1}{2}$, $Q(\psi)$ is just the power dissipated by the circuit when the boundary voltage is ψ, thanks to the principle of minimum power. The factor of $\frac{1}{2}$ simplifies the next proposition, which uses Q to compute the boundary current as a function of the boundary voltage. We will later see that in fact $Q(\psi)$ is a nonnegative quadratic form on $\mathbb{R}^{\partial N}$.

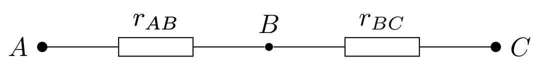
Since Q is a smooth real-valued function on $\mathbb{R}^{\partial N}$, its differential dQ at any given point $\psi \in \mathbb{R}^{\partial N}$ defines an element of the dual space $(\mathbb{R}^{\partial N})^*$, which we denote by dQ_ψ. In fact, this element is equal to the boundary current ι corresponding to the boundary voltage ψ:

Proposition 3.9. *Suppose $\psi \in \mathbb{R}^{\partial N}$. Suppose ϕ is any extension of ψ minimizing the power. Then $dQ_\psi \in (\mathbb{R}^{\partial N})^* \cong \mathbb{R}^{\partial N}$ gives the boundary current of the current induced by the potential ϕ.*

Note this only depends on Q, which makes no mention of the potentials at nonterminals. This is amazing: the way power depends on boundary potentials completely characterises the way boundary currents depend on boundary potentials.

3.2.5 An example

Example 3.10 (Resistors in series). Resistors are said to be placed in series if they are placed end to end or, more precisely, if they form a path with no self-intersections. It is well known that resistors in series are equivalent to a single resistor with resistance equal to the sum of their resistances. To prove this, consider the following circuit comprising two resistors in series, with input A and output C:



Now, the extended power functional $P : \mathbb{R}^{\{A,B,C\}} \to \mathbb{R}$ for this circuit is

$$P(\phi) = \frac{1}{r_{AB}}\big(\phi(A) - \phi(B)\big)^2 + \frac{1}{r_{BC}}\big(\phi(B) - \phi(C)\big)^2,$$

while the power functional $Q : \mathbb{R}^{\{A,C\}} \to \mathbb{R}$ is given by minimisation over values of $\phi(B) = x$:

$$Q(\psi) = \frac{1}{2}\min_{x\in\mathbb{R}}\left(\frac{1}{r_{AB}}\big(\psi(A) - x\big)^2 + \frac{1}{r_{BC}}\big(x - \psi(C)\big)^2\right).$$

Differentiating with respect to x, we see that this minimum occurs when

$$\frac{1}{r_{AB}}\big(x - \psi(A)\big) + \frac{1}{r_{BC}}\big(x - \psi(C)\big) = 0,$$

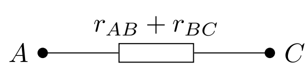
and hence when x is the r-weighted average of $\psi(A)$ and $\psi(C)$:

$$x = \frac{r_{BC}\psi(A) + r_{AB}\psi(C)}{r_{BC} + r_{AB}}.$$

Substituting this value for x into the expression for Q above and simplifying gives

$$Q(\psi) = \frac{1}{r_{AB} + r_{BC}}\big(\psi(A) - \psi(C)\big)^2.$$

This is also the power functional of the circuit



and so the circuits are equivalent.

3.2.6 Dirichlet forms

We have seen that a behaviour is completely specified by the power functional Q on the vector space $\mathbb{R}^{\partial N}$. Now we describe which functionals can arise this way. They are the quadratic forms known as 'Dirichlet forms', and they admit a number of equivalent characterizations. We start with the simplest. A Dirichlet form on S will be a certain sort of quadratic form on $\mathbb{R}^S$:

Definition 3.11. *Given a finite set S, a* **Dirichlet form** *on S is a quadratic form $Q : \mathbb{R}^S \to \mathbb{R}$ given by the formula*

$$Q(\psi) = \sum_{i,j} c_{ij}(\psi_i - \psi_j)^2$$

for some nonnegative real numbers c_{ij}.

Any Dirichlet form is nonnegative: $Q(\psi) \geq 0$ for all $\psi \in \mathbb{R}^S$. However, not all nonnegative quadratic forms are Dirichlet forms. For example, if $S = \{1, 2\}$:

$$Q(\psi) = (\psi_1 + \psi_2)^2$$

is not a Dirichlet form.

It is crucial for us that the property of being a Dirichlet form is preserved under minimising over linear subspaces of the domain generated by subsets of the given basis.

Proposition 3.12. *If $Q : \mathbb{R}^{S+T} \to \mathbb{R}$ is Dirichlet, then*

$$\min_{\nu \in \mathbb{R}^T} Q(-, \nu) : \mathbb{R}^S \to \mathbb{R}$$

is Dirichlet.

This means that all power functionals are Dirichlet forms.

Corollary 3.13. *Let $Q : \mathbb{R}^{\partial N} \to \mathbb{R}$ be the power functional for some circuit. Then Q is a Dirichlet form.*

The converse is also true: simply construct the circuit with set of vertices X and an edge of resistance $\frac{1}{c_{ij}}$ between any $i, j \in X$ such that the term $c_{ij}(\psi_i - \psi_j)$ appears in the Dirichlet form. This gives:

Proposition 3.14. *A function Q is the power functional for some circuit if and only if Q is a Dirichlet form.*

3.2.7 Summary

In summary, in this section we have shown the existence of a surjective function

$$\left\{ \begin{array}{c} \text{circuits with inputs } i(X), \text{ outputs } o(Y), \\ \text{and } i(X) \cap o(Y) = \emptyset \end{array} \right\} \longrightarrow \left\{ \text{Dirichlet forms on } i(X) + o(Y) \right\}$$

mapping two circuits to the same Dirichlet form if and only if they are functionally equivalent.

3.3 Lagrangian relations

Open circuits can be connected with one another to create new circuits, determining a new relationship. We show that this new relationship is a function only of the current-voltage relationships determined by the composite circuits, and hence does not depend on the internal structure of the composite circuits. This allows us to speak of a composition rule for current-voltage relationships.

We begin by constructing a category where the morphisms are, roughly speaking, behaviours. Recall that behaviours, or equivalence classes of circuits, are in one-to-one correspondence with Dirichlet forms. Although one can define a composition rule for behaviours by adding their corresponding Dirichlet forms and minimising over internal nodes, no behaviour acts as the identity for this composition. To construct our category then, we move to a setting which allows more morphisms: that of sympletic vector spaces and their Lagrangian subspaces. These extra morphisms may be interpreted as circuits having wires with zero resistance. As the expression for the extended power functional includes the reciprocals of resistances, such circuits cannot be expressed within the framework we have developed thus far. Indeed, for these idealised circuits there is no function taking boundary potentials to boundary currents: the lack of resistance is thought as as implying that any difference in potentials at the boundary induces 'infinite' currents.

3.3.1 Symplectic vector spaces

Inspired by the principle of least action of classical mechanics in analogy with the principle of minimum power, we turn to symplectic methods to discuss behaviours. A symplectic vector space allows a more invariant perspective, viewing voltages and currents in the same vector space.

A circuit made up of wires of positive resistance defines a function from boundary potentials to boundary currents. A wire of zero resistance, however, does not define a function: the principle of minimum power is obeyed as long as the potentials at the two ends of the wire are equal. More generally, we may thus think of circuits as specifying a set of allowed voltage-current pairs, or as a relation between boundary potentials and boundary currents. This set forms what is called a Lagrangian subspace, and is given by the graph of the differential of the power functional. More generally, Lagrangian submanifolds graph derivatives of smooth functions: they describe the point evaluated and the tangent to that point within the same space.

Definition 3.15. *Given a vector space V over a field $\mathbb{F}$, a* **symplectic form** $\omega : V \times V \to \mathbb{F}$ *on V is a antisymmetric nondegenerate bilinear form. That is, a symplectic form ω is a function $V \times V \to \mathbb{F}$ that obeys*

(i) bilinearity: for all $\lambda \in \mathbb{F}$ and all $u, v \in V$ we have $\omega(\lambda u, v) = \omega(u, \lambda v) = \lambda\omega(u, v)$;

(ii) total isotropy: for all $v \in V$ we have $\omega(v, v) = 0$; and

(iii) nondegeneracy: given $v \in V$, $\omega(u, v) = 0$ for all $u \in V$ if and only if $u = 0$.

A **symplectic vector space** (V, ω) *is a vector space V equipped with a symplectic form ω.*

Given symplectic vector spaces $(V_1, \omega_1), (V_2, \omega_2)$, a **symplectic map** *is a linear map $f : (V_1, \omega_1) \to (V_2, \omega_2)$ such that $\omega_2(f(u), f(v)) = \omega_1(u, v)$ for all $u, v \in V_1$. A* **symplectomorphism** *is a symplectic map that is also an isomorphism.*

A **symplectic basis** for a symplectic vector space (V, ω) is a basis $\{p_1, \dots, p_n, q_1, \dots, q_n\}$ such that $\omega(p_i, p_j) = \omega(q_i, q_j) = 0$ for all $1 \le i, j \le n$, and $\omega(p_i, q_j) = \delta_{ij}$ for all $1 \le i, j \le n$, where δ_{ij} is the Dirac delta, equal to 1 when $i = j$, and 0 otherwise. A symplectomorphism maps symplectic bases to symplectic bases, and conversely: any map that takes a symplectic basis to another symplectic basis is a symplectomorphism.

Example 3.16 (The symplectic vector space generated by a finite set). Given a finite set S, we consider the vector space $\mathbb{R}^S \oplus \mathbb{R}^{S^*}$ a symplectic vector space, with symplectic form

$$\omega((v, \phi), (v', \phi')) = \phi'(v) - \phi(v').$$

Let $\{v_s\}_{s \in S}$ be the basis of $\mathbb{R}^S$ consisting of the functions $S \to \mathbb{R}$ mapping s to 1 and all other elements of s to 0, and let $\{\phi_s\}_{s \in S} \subseteq \mathbb{R}^{S^*}$ be the dual basis. Then $\{(v_s, 0), (0, \phi_s)\}_{s \in S}$ forms a symplectic basis for $\mathbb{R}^s \oplus \mathbb{R}^{S^*}$.

There are a two common ways we will build symplectic spaces from other symplectic spaces: conjugation and summation. Given a symplectic form ω, we may define its **conjugate** sympletic form $\overline{\omega} = -\omega$, and write the conjugate symplectic space $(V, \overline{\omega})$ as $\overline{V}$. Given two symplectic vector spaces $(U, \nu), (V, \omega)$, we consider the direct sum $U \oplus V$ a sympletic vector space with the symplectic form $\nu + \omega$. This is *not* a product in the category of symplectic vector spaces and symplectic maps.

The symplectic form gives a map from a symplectic space to its dual space. We study these dual elements, and the symplectic form itself, through their kernels. Given a subspace S of V, we define its **complement**

$$S^\circ = \{v \in V \mid \omega(v, s) = 0 \text{ for all } s \in S\}.$$

Given a symplectic vector space $V = U \oplus U^*$, the subspace U has the property of being a maximal subspace such that the symplectic form restricts to a zero form on U. Subspaces with this property are known as Lagrangian subspaces, and may be realised as the image of U under symplectic isometries $V \to V$. Lagrangian subspaces may also be viewed as the subspaces that correspond to possible (point, tangent vector) pairs for quadratic forms on V. They may also mostly be viewed as graphs of 'self-adjoint' maps $U \to U^*$.

Definition 3.17. *Let S be a linear subspace of a symplectic vector space (V, ω). We say that S is* **isotropic** *if $\omega|_{S\times S} = 0$, and that S is* **coisotropic** *if S° is isotropic. A subspace is* **Lagrangian** *if it is both isotropic and coisotropic.*

Lagrangian subspaces are also known as Lagrangian correspondences and canonical relations.

3.3.2 Lagrangian subspaces from Dirichlet forms

We are interested in these symplectic structures as Dirichlet forms on S give rise to Lagrangian subspaces of $\mathbb{R}^S \oplus \mathbb{R}^{S^*}$. In this subsection we show how these subspaces are constructed. More generally, we show that there is a one-to-one correspondence between Lagrangian subspaces and quadratic forms.

Proposition 3.18. *Let S be a finite set. Given a quadratic form Q on $\mathbb{R}^S$, the subspace*

$$L_Q = \{(v, dQ_v) \mid v \in \mathbb{R}^S\} \subseteq \mathbb{R}^S \oplus \mathbb{R}^{S^*},$$

where $dQ_v \in \mathbb{R}^{S^*}$ is the value of differential dQ of Q at $v \in \mathbb{R}^S$, is Lagrangian. Moreover, this construction gives a one-to-one correspondence

$$\left\{\textit{quadratic forms on } \mathbb{R}^S\right\} \longleftrightarrow \left\{\begin{array}{c}\textit{Lagrangian subspaces of } \mathbb{R}^S \oplus \mathbb{R}^{S^*} \\ \textit{with trivial intersection with } \mathbb{R}^S \oplus \{0\} \subseteq \mathbb{R}^S \oplus \mathbb{R}^{S^*}\end{array}\right\}.$$

In particular, every Dirichlet form defines a Lagrangian subspace. We call the Lagrangian subspaces arising from Dirichlet forms in this way **Dirichlet Lagrangian subspaces**.

3.3.3 Lagrangian subspaces as morphisms

Recall that a relation between sets X and Y is a subset R of their product $X \times Y$. Furthermore, given relations $R \subseteq X \times Y$ and $S \subseteq Y \times Z$, there is a composite relation $(S \circ R) \subseteq X \times Z$ given by pairs (x, z) such that there exists $y \in Y$ with $(x, y) \in R$ and $(y, z) \in S$—a direct generalisation of function composition. A Lagrangian relation between symplectic vector spaces V_1 and V_2 is a relation between V_1 and V_2 that forms a Lagrangian subspace of the symplectic vector space $\overline{V_1} \oplus V_2$. This allows us to think of certain Lagrangian subspaces as morphisms, giving a way to compose them.

Definition 3.19. *A* **Lagrangian relation** *$L : V_1 \to V_2$ is a Lagrangian subspace L of $\overline{V_1} \oplus V_2$.*

The composite of two Lagrangian relations is again a Lagrangian relation.

Proposition 3.20. *Let $L : V_1 \to V_2$ and $L' : V_2 \to V_3$ be Lagrangian relations. Then their composite relation $L' \circ L$ is a Lagrangian relation $V_1 \to V_3$.*

This composition agrees with our briefly mentioned composition of Dirichlet forms, and hence also composition of circuits.

3.3.4 The dagger compact category of Lagrangian relations

Lagrangian relations solve the identity problems we had with Dirichlet forms: given a symplectic vector space V, the Lagrangian relation $\mathrm{id}: V \to V$ specified by the Lagrangian subspace

$$\mathrm{id} = \{(v,v) \mid v \in V\} \subseteq \overline{V} \oplus V,$$

acts as an identity for composition of relations. We thus have a category.

Definition 3.21. *We write* **LagrRel** *for the category with objects symplectic vector spaces and morphisms Lagrangian relations.*

In fact the move to the setting of Lagrangian relations, rather than Dirichlet forms, adds far richer structure than just identity morphisms. The category LagrRel can be viewed as endowed with the structure of a dagger compact category. We lay this out in steps.

Symmetric monoidal structure

We define the monoidal product of two objects of LagrRel to be their direct sum, and similarly define the monoidal product of two morphisms to be the direct product—note that the direct product of two Lagrangian subspaces is again Lagrangian in the product of their ambient spaces. Defining for all objects U, V, W unitors

$$\lambda_V = \{(0,v,v)\} \subseteq \overline{\{0\} \oplus V} \oplus V,$$
$$\rho_V = \{(v,0,v)\} \subseteq \overline{V \oplus \{0\}} \oplus V,$$

associators

$$\alpha_{U,V,W} = \{(u,v,w,u,v,w)\} \subseteq \overline{(U \oplus V) \oplus W} \oplus U \oplus (V \oplus W),$$

and swaps

$$\sigma_{U,V} = \{(u,v,v,u) \mid u \in U, v \in V\} \subseteq \overline{U \oplus V} \oplus V \oplus U,$$

we have a symmetric monoidal category. Indeed, note that all these structure maps come from symplectomorphisms between the domain and codomain. From this viewpoint it is immediate that all the necessary diagrams commute, and so we have a symmetric monoidal category.

Compactness structure

Each object V of LagrRel is dual to its conjugate space $\overline{V}$, with cup $\eta: \{0\} \to \overline{V} \oplus V$ given by

$$\eta = \{(0,v,v) \mid v \in V\} \subseteq \overline{\{0\}} \oplus \overline{V} \oplus V$$

and cap $\epsilon: \overline{V} \oplus V \to \{0\}$ given by

$$\epsilon = \{(v,v,0) \mid v \in V\} \subseteq \overline{\overline{V} \oplus V} \oplus \{0\}.$$

Dagger structure

Given symplectic vector spaces V_1, V_2, observe that the map

$$\dagger : \overline{V_1} \oplus V_2 \longrightarrow \overline{V_2} \oplus V_1;$$
$$(v_1, v_2) \longmapsto (v_2, v_1)$$

takes Lagrangian subspaces of the domain to Lagrangian subspaces of the codomain. Thus we can view it as a map $\dagger$ taking morphisms $L: V_1 \to V_2$ of LagrRel to morphisms $\dagger(L): V_2 \to V_1$. This defines a functor which is the identity on objects, and in fact a symmetric monoidal dagger functor: by inspection the dagger of the monoidal product of maps is equal to the monoidal product of their daggers, and it takes each structural morphism to its inverse.

This dagger functor plays nicely with the compactness: it is clear that $\eta^\dagger = \epsilon \circ \sigma$. We thus have a dagger compact category.

3.4 The black box functor

We have now developed enough machinery to discuss the semantics of diagrams of resistive circuits and their composition. To recap, we have so far developed two categories: one capturing the notion of the physical form of circuits of linear resistors—in which the morphisms represent these circuits up to topological equivalence—and one capturing the notion of the function of circuits of linear resistors—which contains morphisms representing circuits up to functional equivalence. We now define a functor that maps the physical form of a circuit to its behaviour. In particular, the composition rule for circuits by function reflects the composition rule we use to define circuits by their form. As such a functor identifies all circuits performing the same role, making the internal physical structure of the circuit inaccessible, we call this functor the **black box functor**, writing

$$\blacksquare : \mathrm{ResCirc} \to \mathrm{LagrRel}.$$

On objects this functor maps finite sets to the symplectic vector space generated by this set. That is, let X be an object of ResCirc. Then X is a $(0,\infty)$-graph with no edges or, equivalently, a finite set. We define

$$\blacksquare(X) = \mathbb{R}^X \oplus (\mathbb{R}^X)^*.$$

Defining the way the black box functor acts on morphisms is a bit more involved. Let $\Gamma : X \to Y$ be a circuit. Recall that this means that X and Y are finite sets considered as $(0,\infty)$-graphs with no edges, and Γ is a $(0,\infty)$-graph (N,E,s,t,r) equipped with maps of circuits $i : X \to \Gamma$ and $o : Y \to \Gamma$:

$$\begin{tikzcd} & \Gamma & \\ X \arrow[ur, "i"] & & Y. \arrow[ul, "o"'] \end{tikzcd}$$

To define the image of Γ under our functor $\blacksquare$, we must specify a Lagrangian subspace

$$\blacksquare(\Gamma) \subseteq \overline{\mathbb{R}^X \oplus (\mathbb{R}^X)^*} \oplus \mathbb{R}^Y \oplus (\mathbb{R}^Y)^*.$$

Recall that we associate a Dirichlet form, the power functional $Q : \mathbb{R}^{\partial N} \to \mathbb{R}$ to Γ, where $\partial N = i(X) \cup o(Y) \subseteq N$ is the set of terminals of Γ, and associate a Lagrangian subspace

$$L_Q \subseteq \mathbb{R}^{\partial N} \oplus (\mathbb{R}^{\partial N})^*$$

to this Dirichlet form. It remains to construct from this Lagrangian subspace of $\mathbb{R}^{\partial N} \oplus (\mathbb{R}^{\partial N})^*$ a Lagrangian subspace of $\overline{\mathbb{R}^X \oplus (\mathbb{R}^X)^*} \oplus \mathbb{R}^Y \oplus (\mathbb{R}^Y)^*$.

We make this construction by composing with a relation $\mathbb{R}^{\partial N} \oplus (\mathbb{R}^{\partial N})^* \to \overline{\mathbb{R}^X \oplus (\mathbb{R}^X)^*} \oplus \mathbb{R}^Y \oplus (\mathbb{R}^Y)^*$. Note that an ∂N is the union of the images of $i : X \to N$ and $o : Y \to N$, we may consider these both as maps with codomain ∂N. The induce the dual maps

$$i^* : \mathbb{R}^{\partial N} \longrightarrow \mathbb{R}^X \qquad \text{and} \qquad o^* : \mathbb{R}^{\partial N} \longrightarrow \mathbb{R}^Y$$

given by precomposition of elements of $\mathbb{R}^{\partial N}$ with i and o respectively, and these in turn induce dual maps

$$i^{**} : (\mathbb{R}^X)^* \longrightarrow (\mathbb{R}^{\partial N})^* \qquad \text{and} \qquad o^{**} : (\mathbb{R}^Y)^* \longrightarrow (\mathbb{R}^{\partial N})^*.$$

We then define the relation

$$\begin{aligned} R : \mathbb{R}^{\partial N} \oplus (\mathbb{R}^{\partial N})^* &\longrightarrow \overline{\mathbb{R}^X \oplus (\mathbb{R}^X)^*} \oplus \mathbb{R}^Y \oplus (\mathbb{R}^Y)^* \\ (\phi, \iota) &\longmapsto \{(x, x^*, y, y^*) \mid i^*(\phi) = x, i^{**}(x^*) = \iota, o^*(\phi) = y, o^{**}(y^*) = \iota\}. \end{aligned}$$

To understand this construction more concretely, observe that given an element $\phi \in \mathbb{R}^{\partial N}$, $i^*(\phi) \in \mathbb{R}^X$ maps $x \in X$ to $\phi(i(x))$ and $o^*(\phi)$ maps $y \in Y$ to $\phi(o(y))$, so the relation assigns potentials to the elements of X and Y according to the potential at the terminal of Γ that they map to. On the other hand, given $\iota' \in (\mathbb{R}^{X+Y})^*$, $\tau^*(\iota')$ is the element of $(\mathbb{R}^{\partial N})^*$ which assigns to any terminal $n \in \partial N$ the difference $\sum_{i(x)=n} \iota'(x) - \sum_{o(y)=n} \iota'(y)$ between the pointwise currents through the input and output connections that are mapped to n. This means that the preimage of some $\iota \in (\mathbb{R}^{\partial N})^*$ under τ^* consists of all the ways $\iota' \in (\mathbb{R}^{X+Y})^*$ to apportion the current at each terminal $n \in \partial N$ among the connections that map to it under the input and output maps so that the net current into it is the current at n.

As isomorphisms of cospans of $(0,\infty)$-graphs amount to no more than a relabelling of nodes and edges, this construction is independent of the cospan chosen as representative of the isomorphism class of cospans forming the circuit.

It can be checked that this functor is indeed well-defined. We have the following propositions.

Proposition 3.22. *The black box functor* $\blacksquare : \mathrm{ResCirc} \to \mathrm{LagrRel}$ *respects the dagger structures. That is,*

$$\dagger \circ \blacksquare = \blacksquare \circ \dagger.$$

Proposition 3.23. *The image of the homset* $\hom(X,Y)$ *of* ResCirc *under the black box functor* $\blacksquare$ *is the closure of the* $\overline{\textit{set of Dirichlet}}$ *Lagrangian subspaces in the Lagrangian Grassmanian of the sympectic vector space* $\overline{\mathbb{R}^X \oplus (\mathbb{R}^X)^*} \oplus \mathbb{R}^Y \oplus (\mathbb{R}^Y)^*$.

4 Further lines of investigation

4.1 Networks of resistors

The material developed above still bears investigation. Can we describe the structure of the categories ResCirc and LagrRel better? What are the preimages of morphisms in LagrRel under $\blacksquare$—or, equivalently, how do we tell when two circuits have the same behaviour? The category LagrRel has a Frobenius algebra corresponding to 'networks of wires of zero resistance'—can this be used for anything? What list of properties uniquely characterise the black box functors as a functor ResCirc $\to$ LagrRel?

4.2 General circuits

The work of the previous section restricts attention to those circuits made of linear resistors where the voltages and currents do not depend on time. Although this setting is well-understood without the machinery employed here, generalisation to include other circuit components opens classification problems that are yet to answered. It is anticipated that much of the framework should apply without great change by working over the complex numbers instead of the reals—this allows discussion impedance rather than resistance, and hence include inductors and capacitors as well as resistors.

For example, while in this setting it was proved by Brune that for finite two terminal circuits a function can be realised as an impedance function if and only if it is a rational function with positive real part, analytic on the right-half plane, and takes real values on the real axis [10]. Although partial results generalise this to multiple-input/multiple-output circuits, a full classification is not known. Furthermore, many more open questions arise if constraints are placed on the number of circuit elements.

4.3 Signal-flow diagrams

Baez and Erbele [5] have given a definition of a monoidal category of signal-flow diagrams. What is the relationship between this category, and the categories of circuits and behaviours discussed here?

4.4 Hodge theory

Well-known and dating back to Bott, there is a large body of work on the use of homological and Hodge theoretic methods in understanding circuits of linear resistors, and in particular towards understanding duality between current and voltage [11, 32, 43]. It may be worthwhile to look at more general circuits from this perspective, and integrate these ideas into the monoidal setting developed here.

4.5 Higher categorical concerns

Although we defined ResCirc as a bicategory of cospans, in the above we decategorified immediately and worked in the 1-categorical setting to discuss the semantics. In the higher categorical setting, 2-morphisms between circuits correspond to rewrite rules or graph transformations. These have been studied by Verdiére and collaborators [44, 45], and more recently by Alman, Lian, and Tran [2], and are useful for discussing normal forms of circuits. Here there are also more foundational questions concerning higher categories of spans and cospans.

It would also be interesting to relate this work to that of Spivak [40] on circuits and operads.

Acknowledgements

I thank John Baez, Bob Coecke, Jamie Vicary, and Samson Abramsky for their support and useful discussions over the past year. Figures 1–4 are taken from Wikimedia Commons, used under a Creative Commons licence.

References

[1] S. Abramsky and B. Coecke, A categorical semantics of quantum protocols, in *Proceedings of the 19th IEEE conference on Logic in Computer Science (LiCS04)*, IEEE Computer Science Press, 2004. Also available as arXiv:quant-ph/0402130.

[2] J. Alman, C. Lian, B. Tran, Circular planar electrical networks I: the electrical poset EP_n. Available as arXiv:1309.2697.

[3] J. C. Baez, Quantum quandaries: a category-theoretic perspective, in *Structural Foundations of Quantum Gravity*, eds. Steven French, Dean Rickles and Juha Saatsi, Oxford U. Press, 2006, pp. 240–265.

[4] J. C. Baez, Circuit theory I, draft article, ncatlab. Available at http://ncatlab.org/johnbaez/show/Circuit+theory+I.

[5] J. C. Baez, J. Erbele, Categories in control, in preparation.

[6] P. Bamberg and S. Sternberg, *A Course of Mathematics for Students of Physics 2*, Chapter 12: The theory of electrical circuits, Cambridge University, Cambridge, 1982.

[7] J. Bénabou. Introduction to bicategories I. In *Reports of the Midwest Category Seminar*, Lecture Notes in Mathematics 47, pages 1-77. Springer, 1967.

[8] S. Benenti, *Hamiltonian Structures and Generating Families*, Universitext, Springer, 2011.

[9] J. F. Broenink, Introduction to Physical Systems Modeling with Bond Graphs, in *The SiE Whitebook on Simulation Methodologies*, 1999.

[10] O. Brune, Synthesis of a finite two-terminal network whose driving-point impedance is a prescribed function of frequency, Doctoral thesis, MIT, 1931. Available at http://dspace.mit.edu/bitstream/handle/1721.1/10661/36311006.pdf?sequence=1.

[11] G. E. Ching, Topological concepts in networks; an application of homology theory to network analysis, *Proc. 11th. Midwest Conference on Circuit Theory*, University of Notre Dame, 1968, pp. 165-175.

[12] J. Clerk Maxwell, On governors, *Proceedings of the Royal Society* **100** (1868). Available at http://rspl.royalsocietypublishing.org/content/16/270.full.pdf

[13] D. Cimasoni and V. Turaev, A Lagrangian Representation of Tangles, *Topology* **44** (2005), 747-767.

[14] B. Friedland, *Control System Design: An Introduction to State-Space Methods*, McGraw-Hill, 1986.

[15] M. Fukushima, *Dirichlet Forms and Markov Processes*, North-Holland, Amsterdam, 1980.

[16] P. W. Gross and P. R. Kotiuga, *Electromagnetic Theory and Computation: A Topological Approach*, Cambridge University Press, 2004.

[17] J. H. Hubbard, The Bott–Duffin synthesis of electrical circuits, in *A Celebration of the Mathematical Legacy of Raoul Bott* ed. P. R. Kotiuga, CRM Proceedings and Lecture Notes, American Mathematical Society, 2010.

[18] A. Joyal and Ross Street, The geometry of tensor calculus I, *Adv. Math.* **88** (1991), 55113.

[19] A. Joyal and Ross Street, The geometry of tensor calculus II. Available at http://www.math.mq.edu.au/~street/.

[20] H. W. Knobloch and M. Thoma, Aspects of Modern Control Theory, *Perspectives in Mathematics: Anniversary of Oberwolfach 1984*, Birkhäuser Verlag, Basel, 1984, pp. 337-376.

[21] P. Katis, N. Sabadini, R. F. C. Walters, On the algebra of systems with feedback and boundary, *Rendiconti del Circolo Matematico di Palermo Serie II, Suppl.* **63** (2000), 123-156.

[22] J. Kigami, *Analysis on Fractals*, Cambridge U. Press. First 60 pages available at http://www-an.acs.i.kyoto-u.ac.jp/~kigami/AOF.pdf.

[23] Z.-M. Ma and M. Röckner, *Introduction to the Theory of (Non-Symmetric) Dirichlet Forms*, Springer, Berlin, 1991.

[24] S. Mac Lane, Natural associativity and commutativity, *Rice Univ. Studies* **49** (1963), 2846.

[25] S. Mac Lane, *Categories for the Working Mathematician*, Springer, 1998.

[26] G. Ohm, *Die Galvanische Kette, Mathematisch Bearbeitet*, T. H. Riemann, Berlin, 1827. Also available at www.ohm-hochschule.de/bib/textarchiv/Ohm.Die_galvanische_Kette.pdf.

[27] H. T. Odum, *Ecological and General Systems: An Introduction to Systems Ecology*, Colorado University Press, 1994.

[28] H. F. Olson, *Dynamical Analogies*, 2nd ed., Van Nostrand, 1958.

[29] J. W. Polderman, J. Willems, *Introduction to the Mathematical Theory of Systems and Control*, Springer, 1998.

[30] H. M. Paynter, *Analysis and Design of Engineering Systems*, MIT Press, Cambridge, Massachusetts, 1961.

[31] R. Penrose. Applications of negative dimensional tensors. In *Combinatorial Mathematics and its Applications*, ed. D. J. A. Welsh, Academic Press, New York, 1971, pp 221–244.

[32] J. P. Roth, Existence and uniqueness of solutions to electrical network problems via homology sequences, *Mathematical Aspects of Electrical Network Theory*, SIAM-AMS Proceedings III, 1971, pp. 113-118.

[33] C. Sabot, Existence and uniqueness of diffusions on finitely ramified self-similar fractals, Section 1: Dirichlet forms on finite sets and electrical networks, *Annales Scientifiques de l'cole Normale Suprieure*, Sr. 4, 30 (1997), 605-673. Also available at http://www.numdam.org/numdam-bin/item?id=ASENS_1997_4_30_5_605_0.

[34] C. Sabot, Electrical networks, symplectic reductions, and application to the renormalization map of self-similar lattices, *Proc. Sympos. Pure Math.* **72** (2004), 155-205. Also available as arXiv:math-ph/0304015.

[35] P. Selinger, Dagger compact closed categories and completely positive maps, in *Proceedings of the 3rd International Workshop on Quantum Programming Languages* (QPL 2005), *ENTCS* **170** (2007), 139163. Also available at http://www.mscs.dal.ca/~selinger/papers/dagger.pdf.

[36] P. Slepian, *Mathematical Foundations of Network Analysis*, Springer, Berlin, 1968.

[37] S. Smale, On the mathematical foundations of electrical network theory, *J. Diff. Geom.* **7** (1972), 193-210.

[38] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Texts in Applied Mathematics 6, Springer 1998. Available at http://www.math.rutgers.edu/~sontag/FTP_DIR/sontag_mathematical_control_theory_springer98.pdf.

[39] E. D. Sontag, Some new directions in control theory inspired by systems biology, *Syst. Biol.* **1**(1) (2004), 9-18. Available at http://popmodel.narod.ru/sb/9.pdf.

[40] D. I. Spivak, The operad of wiring diagrams: Formalizing a graphical language for databases, recursion, and plug-and-play circuits, 2013. Available as arXiv:1305.0297.

[41] M. Stay, Compact Closed Bicategories, 2013. Available as arXiv:1301.1053.

[42] J. U. Thoma, *Introduction to Bond Graphs and Their Applications*, Pergamon Press, Oxford, 1975.

[43] Y. C. de Verdiere, Théorème de Kirchhoff et théorie de Hodge, *Séminaire de Théorie spectrale et géométrie* **9** (1990-1991), 89-94.

[44] Y. C. de Verdiere, Reseaux electriques planaires I, *Comment. Math. Helv.* **69** (1994), 351-374. Also available at http://www-fourier.ujf-grenoble.fr/~ycolver/All-Articles/94a.pdf.

[45] Y. Colin de Verdiere, I. Gitler, and D. Vertigan, Réseaux électriques planaire II, *Comment. Math. Helv.* **71** (1996), 144167.

[46] A. Weinstein, Symplectic geometry, *Bull. Amer. Math. Soc.* **5**(1) (1981), 1-13.

[47] J. Willems, The behavioral approach to open and interconnected systems, *IEEE Control Systems Magazine* **27** (2007), 46-99. Also available at http://homes.esat.kuleuven.be/~jwillems/.

[48] J. Willems, In control, almost from the beginning until the day after tomorrow, *European Journal of Control* **13** (2007), 71-81.