

The Network

Dan Auerbach

2015

Part I

The Telephone

In 1876, to celebrate the centennial of the Declaration of Independence, the United States hosted its first World's Fair in Philadelphia. The event lasted several months, thirty-seven countries participated, and nearly 10 million visitors attended.

The biggest draws could safely be said to be mechanical contraptions: displays of steam engines, the world's Remington Typographic Machine (typewriter), and precision watches, to name a few. Yet the age of electricity was just around the corner, and electrical inventions were also on display. Tucked away in a small area across the Machinery Hall sat a young Alexander Graham Bell, demonstrating the transfer of human voice through wires.

People viewed early telephone prototypes as a kind of toy, not a serious technological contender to the business-oriented telegraph empire of Western Union. That perception was not just short-sightedness; these prototypes were limited in their range, typically connecting just two adjacent rooms, and it was not obvious at first that the technology could be improved upon to rival the telegraph network. Yet once it became clear that longer-distance telephony was possible, it did not take long for commercial telephony to develop, beginning a relatively rapid shift in power away from the well-established Western Union and instead towards AT&T, then a small start-up corporation that would come to define the telecommunications industry in the United States for a century.

In this part of the book, we will examine issues related to how the telephone works: what is sound, how is an analog signal transmitted over a wire, what infrastructure allowed AT&T to provide telephone access to the vast majority of Americans in their home, and how the regulatory landscape began to evolve as the telecommunications industry matured.

Chapter 1

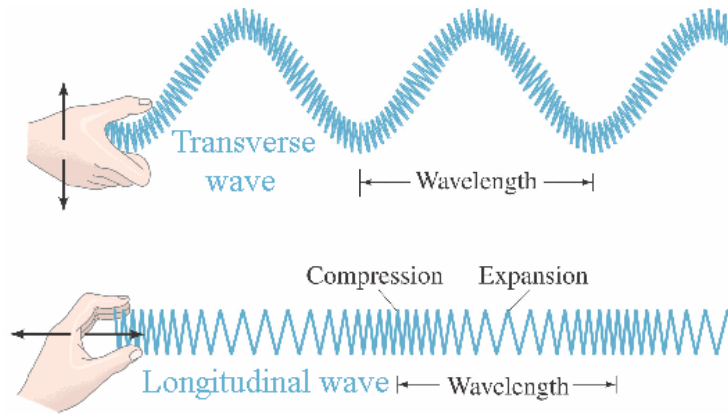
Signals

1.1 Sound

When a tree falls and hits the ground, some of the energy upon impact is transferred to nearby air particles, pushing them outwards in a wave.



The wave is a *longitudinal* wave, meaning it does not have the peaks and troughs of an ocean wave perpendicular to the direction of travel, but instead expands and contracts along a single dimension as it travels.



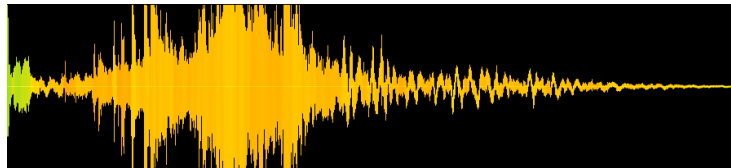
Copyright © 2005 Pearson Prentice Hall, Inc.

Imagine a nuclear explosion, except that instead of an outwardly expanding sphere of fire, there is an outwardly expanding sphere where air particles are more tightly bunched together for a split second. This momentary compression of air particles constitutes a change in *air pressure* in that location.

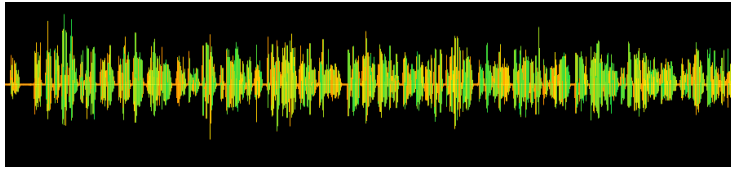
As objects on Earth interact with each other, waves of pressure expand outward in the surrounding medium (in the example above, this medium is the air). These pressure changes often signal events that are important to an organism, like when something might eat you or when you might be able to eat something. Given the importance of this information for survival, many species including humans evolved auditory systems to make sense of the constantly and subtly changing air pressure all around us; we call our experience of this phenomena *hearing*. And in parallel to the development of our sense of hearing, we evolved increasingly sophisticated ways to create sound ourselves, culminating in vocal chords and speaking.

The simplified description above of a tree falling might lead one who is unfamiliar with hearing to believe that there is only a single and uniform experience of a sound, but of course that's not the case. Sound has many different qualities, like pitch, intensity, duration, and timbre. It is quite easy to tell the singing of a friend from the sound of a tree hitting the ground. This differentiation is possible because sound waves can be rather sophisticated, the miniscule air particles compressing and un-compressing in non-uniform ways over very small distance and time scales.

To highlight these differences, here is a representation of the sound of a tree falling:



And here is a representation of human speech:

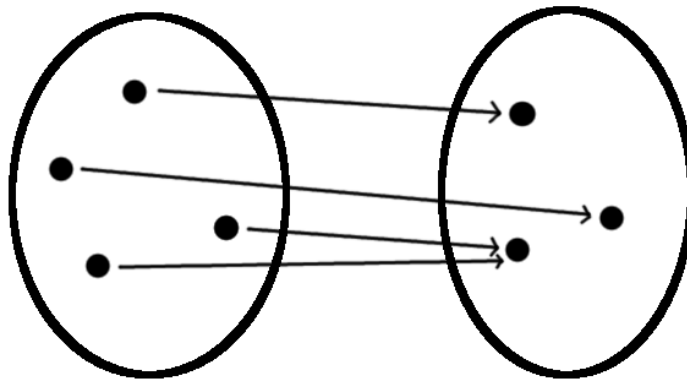


The above figures are called *time domain representations*: the x axis is time, and the y axis represents the instantaneous change in air pressure. Large values on the y axis mean that particles are tightly compressed at that instant. Of course these representations have to be taken from a single source point: two people standing in different places may have different auditory experiences of the same tree-falling event.

1.2 Functions

The two figures above representing various sounds are examples of *functions*, a central concept in mathematics.

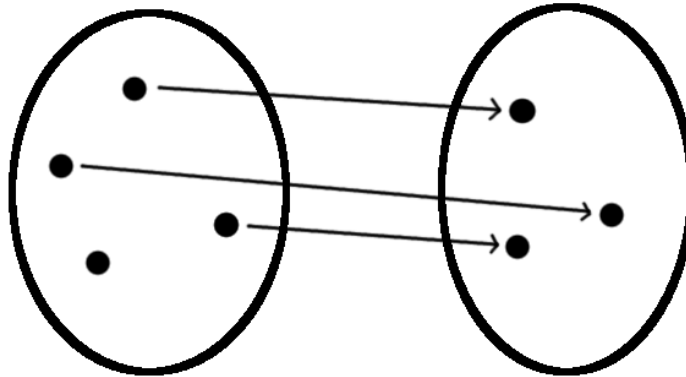
A function is a mapping from one set of objects to another:



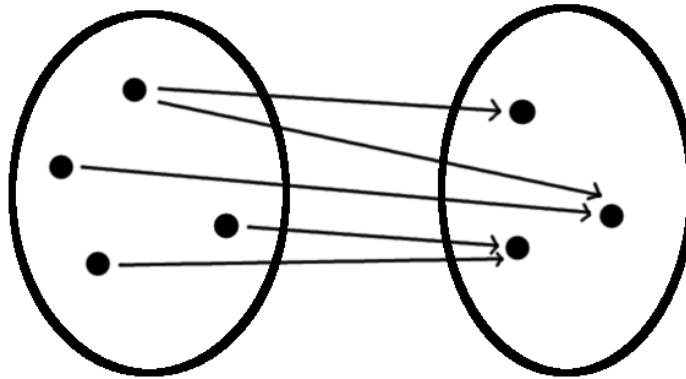
We say that a function has a *domain*, which is the set of objects on the left in the figure above, and a *co-domain* which is the set of objects on the right. There is one important rule that must be followed for something to count as a function:

- For each object in the domain, a function **MUST** map that to **EXACTLY ONE** object in the co-domain

In other words, since functions must be defined on the entire domain, this isn't allowed:



And neither is this:



Note that the converse of the second rule above does not hold; it is perfectly acceptable for several values in the domain to be mapped to a single value in the co-domain, as is the case in the first diagram above.

So far, the domain has been 4 blobs, and the co-domain has been 3 blobs, but this is just one example. There are no restrictions on the domain or co-domain. You can have a function where the domain is pet owners on Earth, and the co-domain is the set of pets:

[IMAGE]

We might call this function “FavoritePet”, and write $\text{FavoritePet}(\text{John}) = \text{Fluffy}$ to mean that the function “FavoritePet” maps the element “John” in the domain of pet owners to the element “Fluffy” in the domain of pets, a formal way of writing that Fluffy is John’s favorite pet. (Exercise: would “Owns” with the same domain and co-domain constitute a function?)

To indicate that a function f has domain A and co-domain B , we use the notation $f : A \rightarrow B$.

The domain and co-domain of a function can be finite (like Pet Owners) or infinite (like natural numbers: $\mathbb{N} = 0, 1, 2, 3, \text{etc.}$). The domain and co-domain could be the same, but they do not have to be.

Sometimes functions can be described easily, like “FavoritePet” or the function $f : \mathbb{N} \rightarrow \mathbb{N}$ that takes a natural number and doubles it. We could call this function “Doubles”, but mathematicians prefer a more compact and precise description: $f(x) = 2x$. Other times functions cannot be described easily at all: they are simply a random looking mapping from the domain to the co-domain. In this case, the only way to give a full description of the function is to give its value for every single input object, a task that is sometimes provably impossible when there are an infinite number of input objects.

Domains or co-domains might be multi-dimensional spaces. For example, a function $f : \mathbb{N}^3 \rightarrow \mathbb{N}$ maps triplets like $(4, 5, 1)$ to natural numbers like 17774. Domains and co-domains can also be discrete, like the natural numbers. Or they can be continuous, like pitch, or time.

A function is an abstract object, and as such it has many representations. For example, the function $f(x) = 2x$ can also be described by plotting it on a two dimensional axis, where the x-axis is the input of the function and the y-axis is the output:

[PLOT OF $f(x)=2x$]

It could also be represented with the primary visualization that we used to introduce the notion of a function:

[IMAGE OF DOTS]

We will return to functions all the time, so it is important to develop a sharp understanding of functions.

Exercise: write out a mathematical description for a function f that takes two numbers, and raises the first to the power of the second, then subtracts the second. What is the value of $f(3, 4)$? As a bonus, use graphing software to create a graphical representation of this function as a surface.

1.3 Speech on a wire: an analog approach

Above, we saw examples of functions that represented sound. Here again is our time domain representation of speech, also called a speech *waveform*:

[IMAGE of speech]

This is just a function in which the domain is time and the co-domain is air pressure. The information in the figure above is enough to create a sound that we would recognize as speech. In order for telephony to be possible, we need a

way to continuously transmit the information of the sound waveform from one location to another.

An electrical wire does not have air particles that can be compressed and uncompressed, but it does have other properties that can be varied continuously over time. In particular, the current flowing through a wire can increase and decrease. This is how the telegraph worked except that in that case, the current would be on or off, resulting in a function that might look like this:

[IMAGE OF TELEGRAPH CURRENT SIGNAL]

The telephone has to solve the more subtle problem of receiving the human speech form, reproducing the information it contains by varying electrical current, and then reproducing the sound at the other end:

[CHART OF CONCEPTUAL OUTLINE OF TELEPHONE]

1.4 Basics of signals

We will see how the telephone accomplishes this shortly, but first let us turn our attention to the idea of a *signal*.

We have already covered symbolic systems, and the fact that any symbolic representation of information is interchangeable [ADD NAME FOR THIS THEOREM?]. But what happens when the information we care about is something like the human voice that cannot be easily represented as a set of symbols? Recall that in this scenario, we say that the information is *analog*.

There is no single, neat definition of a signal. For our purposes, however, let us use the following slightly narrow and restrictive definition:

[TODO FIX THIS DEFINITION]

Definition 1 *A signal is a function where the domain is time.*

1.5 Information content and analog signals

How much information does a sound signal contain? It depends quite a bit on the signal. A periodic signal like an alarm, intuitively, contains only information for a single period.

[IAMHERE]

Chapter 2

How A Telephone Works

The telephone transforms a sound signal into an electrical signal and back again. It was not the only device in the late 1800s that was seeking to gain mastery over sound.

The phonograph was invented in 1877 and was the original technology in a line that continued with record players, cassette tapes players, CD players, and MP3 players. [FOOTNOTE: it seems this line of technology may have been subsumed by more general purpose desktop and mobile computers in the last 5 years or so]. The phonograph was aimed at solving a similar problem to the telephone, in that it needed to capture a sound signal somehow. The difference is that once captured, instead of the moving the sound signal quickly over a long distance, the phonograph instead was charged with creating a physical object from which the sound could be later retrieved, a record. Early phonographs used a circular disk as the record, and a sound signal was captured by deviations in a spiral groove. [MORE HERE?]

[DIAGRAM OF TELEPHONE VS PHONOGRAPH]

2.1 A conceptual telephone

A one-way telephone has a microphone at one end capable of translating an analog sound signal into an electrical one, and a receiver at the other end capable of translating the electrical signal back into an audible sound signal:

[DIAGRAM]

2.1.1 Microphone

A *microphone* serves as the input mechanism, translating our sound signal into a signal of electrical currents. The main type of microphone for years was the carbon microphone and it worked like this:

[INSERT IMAGE]

Much like our eardrums, a thin metal diaphragm is used that vibrates in response to sound waves. When it is pushed inward, as in the diagram above, the carbon granules are squeezed closer together, decreasing the resistance of the granules, and thus increasing the current flowing through the wire. This means that the wire now can carry the same signal as the sound waves.

You might imagine that the carbon granules have to be specially placed for this scheme to work, but this is not the case. We do not need to know the exact path that electrons might take for the macroscopic principle to hold: as the instantaneous air pressure increases, the diaphragm is pushed inward, packing the granules more tightly and decreasing their electrical resistance, proportionally to the increase in air pressure.

2.1.2 Receiver

A *receiver* is a diaphragm attached to an electromagnet. As the electrical current goes up, the electromagnet has a stronger magnetic pull, causing the diaphragm to vibrate more. As it vibrates, it generates changes in air pressure which are emanated from the receiver as sound.

2.2 Telephones of the early 1900s

In the above conceptual diagram, the electrical circuit is always closed, meaning that current is flowing continuously through the circuit without any means to stop it (aside from cutting the cord). In practice, a mechanism is needed to open the circuit when it is not in use, which helps to save battery. It is also desirable for phones to work in the familiar two-way fashion in which a person can talk and listen at the same time.

Chapter 3

The Telephone Network

The first telephones connected a pair of adjacent rooms, making the problem of knowing who was at the other end of the line pretty straight forward.

But the dream of industrialist Theodore Vail, president of AT&T from 1885 to 1889 and again from 1907 to 1919, was universal telephone service for all:

[QUOTE FROM VAIL]

He was an unabashed monopolist, and genuinely believed that a telecommunications monopoly was in the best interest of the country.

[QUOTE FROM VAIL]

This created a host of challenges: overcoming competition, getting the government to agree with the monopolistic approach to telephony, and establishing service in the most remote and rural parts of America. But alongside these business related challenges, a couple of technical issues loomed large when it came to building out a telephone network. First, how are people going to be identified? Second, how is person A going to be able to talk to person B?

3.1 Telephone Numbers

Naming was becoming a problem for the telegraph in the late 1800s. Until that point, if a telegram was addressed to John Smith in Boston, the message could properly be routed to the Boston office but then what? A lot of responsibility for getting this right fell on the sender of the message to appropriately specify a recipient (*John Smith of Commonwealth Ave in Back Bay*) and on the local courier boys who worked for the telegraph office to have some familiarity with the neighborhood and where people or businesses were located.

As cities grew and telegram volume grew, this ad hoc system was becoming unsustainable.