19

FREQUENCY CONCEPTS

Frequency has long been recognized as a fundamental part of nature—basic to music, all matters of sound transmission, mechanical vibrations, and the mathematics of differential equations. Centuries ago it was known that the frequency of vibrating strings or of resonant flutes can be controlled by adjusting physical parameters such as tension, length, or volume. But deeper understanding of frequency—its true mastery—developed rapidly with the advent of modern communication technology, beginning with the telegraph.

Indeed, frequency concepts play a major role in almost all forms of electrical and optical communication, but these concepts, although now clear and in many respects simple, were originally elusive and mysterious. They were discovered and honed only through patience, genius, and happenstance, tinged by human frailty and fierce competition. Today the theory of frequency is regarded as both beautiful and powerful, and its myriad applications are enjoyed by many.

The next few chapters study several theoretical concepts that comprise the mastery of frequency. This study emphasizes historical aspects more than do earlier chapters, for this history—of the telegraph, telephone, radio, and modern communication—is a significant part of our culture, and one can learn from the missteps as well as brilliant insights of key characters in this development. Through this history we will witness how frequency was mastered.

Here are five basic principles of frequency:

1. Invariance of sinusoids. Consider a continuous signal that passes through a linear medium such as an electric circuit. As a general rule, the result differs in form from the original. For example, perhaps you have listened to your voice spoken under water, and found it highly distorted, or noticed that if you tap on a musical string, the response is much more sustained than a tap on a piece of wood. Graphs of the entering and exiting signals in a linear system will, in general, differ in shape.

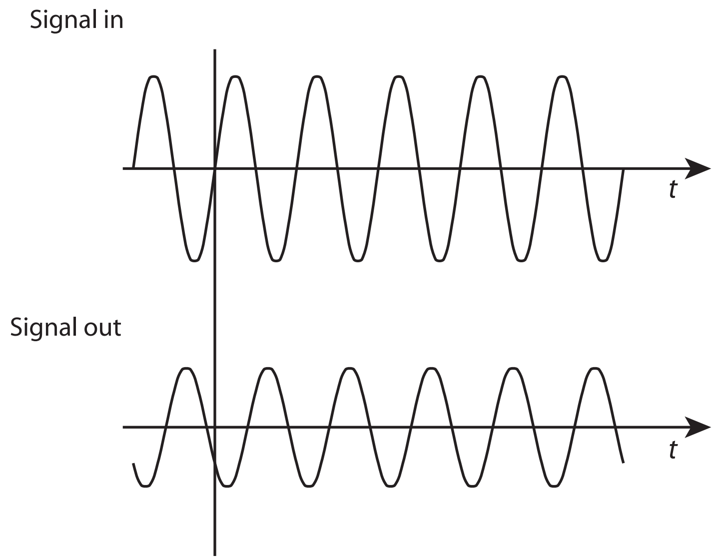


FIGURE 19.1 Invariance of a sine wave. A sinusoidal signal passing through a linear medium or system preserves its sinusoidal character (of the same frequency). Only amplitude and phase are affected.

However, if the original signal is a sinusoid (a time function of the form[1] $A\ \sin(2\pi ft + \theta)$) of frequency f, the resulting output will also be sinusoidal of the same frequency f, although the amplitude and phase may differ from the original (that is, the peaks may be of different height and shifted in time from those of the original). See figure 19.1. The reason that the basic shape is preserved is that a linear system consists of elements that perform some or all of the following operations on a signal: multiplication by a constant, differentiation, integration, delay, and sums of these. From calculus, we know that sine and cosine waves remain sinusoidal of the same frequency under each of these operations.

This invariance of form allows a simple analysis of the behavior of a sinusoidal wave as it passes through a cable, a circuit, or a physical medium such as water or metal. It is not necessary to deduce the output shape, only its amplitude and phase. Much of modern systems analysis is based on this concept.

2. Fourier representation. In the early 1800s the great French mathematician Jean Baptiste Joseph Fourier deduced that any periodic function (within broad limits) can be expressed as a combination of sinusoidal waves. Although he developed this theory as early as 1807, publication was held up by controversy regarding the validity of his methods, and

[1]Here A is a constant, f is the frequency, t is a variable representing time, and θ is a phase shift with $0 \leq \theta \leq 2\pi$. For example, the function $\cos 2\pi t$ is a sinusoid of frequency $f = 1$ with $\theta = \pi/2$.

his work was not publicly available until 1822, when he published his *Théorie analytique de la chaleur*, which not only presented the Fourier series concept but also applied it to solve complicated partial differential equations related to heat transfer. His method was later extended to the Fourier transform that can be used to represent functions that are not periodic.

The fact that an arbitrary function can be represented by a series of sinusoids has profound implications when combined with the principle of invariance discussed above. It means that the linear transformation of an arbitrary signal can be analyzed in three steps: (1) represent the signal as a combination of sinusoids, (2) compute the effect on each of those separate sinusoidal waves using the invariance principle, and (3) combine the separate results to determine the final result. This idea is used, directly or indirectly, in almost all analyses of continuous signals.

3. Resonance. While the previous two concepts are mathematical in nature, the phenomenon of resonance is easily experienced physically. You witness it when you blow across the mouth of a soda bottle to produce a deep tone characteristic of the shape and volume inside, or when you hear a harpist pluck a string. Electrical circuits display resonance as well. Gradually it was learned that it is possible to manipulate resonance to create oscillations, tune radios, and filter out noise.

4. Modulation and the heterodyne principle. Eventually it was discovered that nonlinear transformations of signals, such as squaring or multiplying two signals together, produced important frequency transformations. From elementary trigonometry it is known that when a sine wave of frequency f_1 is multiplied by a sine wave of frequency f_2, the result can be expressed as the sum of two sinusoidal waves—one of frequency $f_1 + f_2$ and the other of frequency $|f_1 - f_2|$. This mathematical identity is manifested physically by the production of a wave that consists of two component frequencies. Purposeful control of this effect is inherent in amplitude modulation (AM) and it also underlies the more general heterodyne principle by which signal frequencies are changed.

5. Nyquist–Shannon sampling theorem. If a continuous signal is sampled at regular intervals, the resulting sequence of sample values can be regarded as an approximation to the actual signal, but obviously some detail may be lost. The Nyquist–Shannon sampling theorem states that if the frequencies contained in the original signal are all below some value W, and the signal is sampled at least every $1/2W$ seconds, then the original signal can be exactly reconstructed from the samples. This result plays a fundamental role in modern communications technology that relies on digital processing. The result also forms the basis of Shannon's most famous capacity calculation, which is presented in chapter 21.

These key principles along with other concepts were discovered gradually along the path to frequency mastery, and we shall follow that development during the course of the next few chapters.

19.1 The Telegraph

The nature of communication technology was fundamentally changed by the invention of the electric telegraph, for it made possible almost instant communication around the world—over continents and across the seas. Messages of national emergency, of family crises, of economic developments, of business relations, and of every aspect of human experience could be sent almost routinely. The world was profoundly different.

The electric telegraph also initiated the modern theory of information and communication. It was the first application of electricity to practical life, and as such it was a visible and friendly breeding ground for fundamental research, the results of which played a significant role in later inventions. Many aspects of electricity were tested and advanced in the course of telegraph development, including the understanding of electrical capacity, induction, batteries, long cables, measurement devices, relays, and so forth.

We study the telegraph and later inventions partly for their interest but mainly because the telegraph motivated a major step in the understanding and use of frequency.

A method for long-distance communication over land was initiated in 1791 when Claude Chappe and his brother transmitted messages visually by sequentially displaying black and white panels that could be observed at a distance of a few miles using a telescope. Their invention was coined *télégraphe*, meaning "far writer." The primitive system was improved by the development of mechanical semaphores between line-of-sight stations. Usually, a large building on a hill was equipped with a tower with long mechanical armlike appendages that could be rotated to various positions to represent alphabetical characters or codewords (figure 19.2). Messages could be sent for several miles, and even further by relaying messages from station to station. These telegraph systems were deployed widely throughout Western Europe and in parts of the United States over the next several decades. Even today their past is evidenced by the existence of a "Telegraph Hill" in many cities.

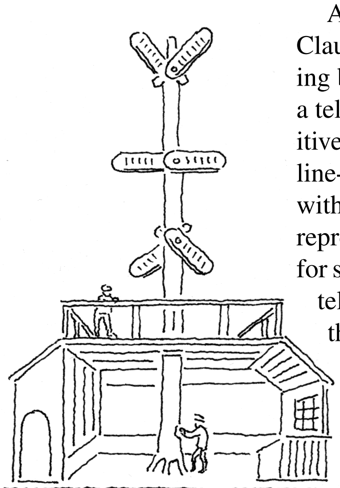


FIGURE 19.2 An early telegraph station. Stations used codes that could be observed from several miles away.

During that period, several individuals imagined an "electric telegraph" using the then newly discovered phenomenon of electricity, generated in those days as static electricity. Real progress was made by Alessandro Volta's invention of the battery, which generated electricity chemically using metals and acid configured into a "pile," and by the revolutionary discovery in 1820 of electromagnetism by the Danish professor Hans Christian Oersted. Oersted observed that a current in a wire influenced the direction of a nearby compass needle, thus establishing a connection between electricity and magnetism. Soon electromagnets and galvanometers were invented that responded to the presence of electric current. Some people attempted to create a telegraph based on these principles, but without significant success.

In 1832 on a return voyage to the United States from Europe, Samuel Morse, a fairly accomplished artist, heard passengers discuss the phenomenon of electromagnetism and the possibility of detecting a current at any point along a wire. He immediately saw that this could lead to the construction of an electric telegraph, and (incorrectly assuming that this idea was original) he determined to develop such a system.

With the help of the physicist Joseph Henry and others, he eventually produced a working telegraph. He also developed the **Morse code** using dots and dashes to correspond directly to letters of the alphabet and numbers. He assigned short code words to frequently used letters (using "dot" for "e," for example), determining the relative frequencies on the basis of a visit to a printer, where he observed the relative numbers of different letters available in a type box.

After four years of difficult negotiations, Morse eventually convinced a skeptical U.S. Congress to support the construction of a 40-mile telegraph line between Baltimore, Maryland, and Washington, D.C. On May 24, 1844, the line operated successfully, reportedly sending "What hath God wrought!" as the first transmission.[2] Over the next decade the telegraph system grew enormously, and its influence was felt in practically every area of life. By 1850 there were over 12,000 miles of telegraph line in the United States.

The main component of the telegraph was an electromagnet, made by wrapping a length of wire many times around an iron core. When current was passed through the coil, a magnetic field was produced. This actuated an armature that indicated the presence of the current.

Electromagnets also form the basis for relays. In this case the armature acts like a switch in a separate circuit: when the first circuit is active and pulls the armature down, the second circuit is completed. By this procedure a relatively weak current in the coil circuit controls what could be a much larger current in the second circuit. Morse understood the importance of the principle that something small could control something large. When his associate Gale worried that the telegraph signal would be too weak at even modest distances, Morse responded, "If a lever can be moved at any distance, it can operate a control point and send a strong signal to the next point, and so on around the globe if desired."

19.2 When Dots Became Dashes

As the telegraph system expanded in both the United States and Europe, it was natural to envision a telegraph crossing bodies of water, including the Atlantic Ocean. Once batteries, relays, and basic line configurations were established, telegraph over land did not seriously challenge the concepts of electric transmission. Over land, voltages could be periodically bumped up along a line by relay stations. And since dots and dashes were coded by hand with telegraph keys and interpreted by ear, the rates of transmission were limited to about 25 words per minute. This meant that dots and dashes were relatively long in duration, as compared to the capabilities of the line. However, transmission through undersea cables was found to be an entirely different matter.

The most obvious difficulties associated with underwater transmission were purely mechanical, but after a few failed attempts, it was found that sufficiently heavy cables insulated with the gummy material extracted from the gutta-percha tree[3] were reasonably reliable. Such cables were laid across the English Channel and other spans of water up to about 100 miles.

However, once the early cables were in place, a fundamental difficulty was discovered. The lines were electrically sluggish, and hence messages were garbled. Dots blurred and became indistinguishable from dashes. Consequently, it was necessary to transmit more slowly than on land-based systems. A typical resulting response to

[2]There is controversy over when this famous phrase was actually first transmitted by telegraph.
[3]This material was used in early golf balls and is currently used in dentistry to fill root canals.

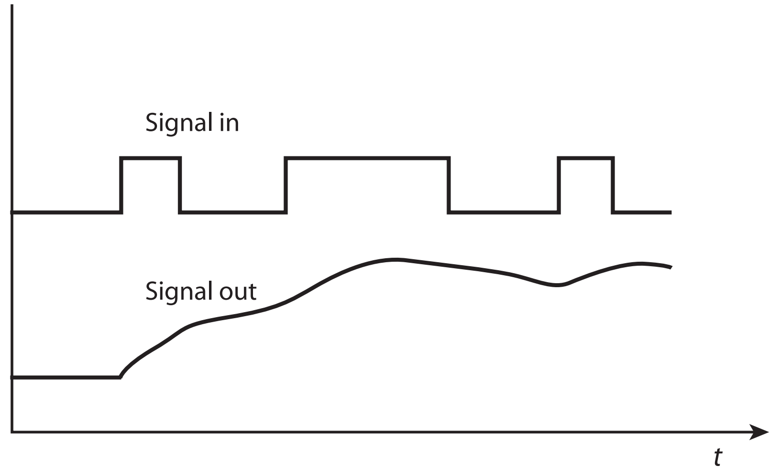


FIGURE 19.3 Response to dots and dashes. When a line has high capacitance, the signals tend to blur together, which necessitates that the rate of signaling be significantly slowed for clear reception.

a series of dots and dashes is shown in figure19.3. The received signal is a smoothed version of the original transmission, making it difficult to distinguish dots from dashes.

Scientists understood that, because the conductor in the cable would be in close proximity to the ocean water, which is also a conductor, a large level of electrical capacitance would be associated with the line. However, it was not understood what the consequences would be.

William Thompson (later Lord Kelvin) was deeply involved in the undersea telegraph project, and developed much of the associated theory. He based his work on the Fourier series, for he was well aware of Fourier's analysis of the heat equation in his famous 1822 publication; after all, Thompson was an expert on heat.[4]

Thompson's analysis explained the effect of undersea line capacity. Capacity is the ability to store charge. Thompson explained the effect of capacity in cables by stating that the cable acted like a thin flexible tube through which water was to be sent. It would be sluggish because the tube would first have to be stretched. Or, as a scientifically more accurate analogy, he explained that sending a pulse along the cable was like sending a pulse of heat along a thin metal wire. The entire wire must be heated before the far end would register a temperature change.

The laying of the Atlantic cable was one of most spectacular and difficult engineering feats of modern times, and it was fraught with a series of disasters. Finally, after several failed attempts at laying a cable, the *Great Eastern*, a huge ship, three times as large as any other ship in service at that time and capable of storing the entire length of nearly 3,000 miles of cable in its hold, was commissioned. On the first attempt with this vessel, in 1865, cable was laid extending about 70 percent across the Atlantic from Ireland to Newfoundland before the cable broke and the venture was

[4]Fourier also was an expert on heat. It is reported that he was in fact obsessed with heat, keeping his rooms unusually warm, perhaps as a result of the three harsh years he spent on a mission to Egypt with Napoleon.

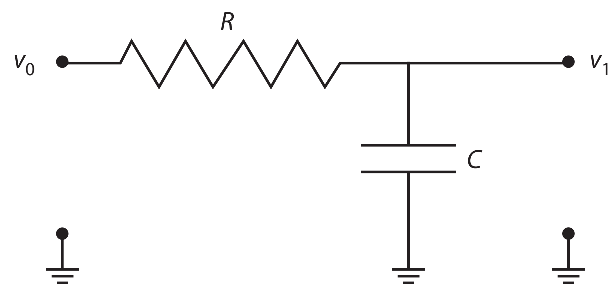


FIGURE 19.4 Highly simplified transmission cable. A transmission cable has resistance R due to the resistance of the wire conductor. There is capacitance because the cable is in close proximity to the sea, which serves as the other side of the circuit. This version is simplified in that it does not reflect the fact that resistance and capacitance are present continuously along the line.

terminated. The second attempt, a year later in 1866 again with the *Great Eastern*, was successful, and another mission was immediately launched that recovered the cable of 1865 and completed it as well. Thompson traveled on these ventures, and provided invaluable technical support based on his theory.

When the two cables were in place, an engineer in Ireland requested that the two ends be connected in Newfoundland. Then from a borrowed silver thimble, some acid, and a piece of zinc he improvised a tiny battery and sent a signal across the Atlantic and back. The returning pulse was detected by a mirror galvanometer after traveling over 4,000 miles.

Thompson's fundamental analysis can be simplified by approximating a telegraph line by the simple circuit shown in figure 19.4, consisting of line resistance R (in **ohms**) due to the resistance of the wire, and capacitance C (in **farads**) resulting from the proximity of the wire to the seawater through the insulation.[5]

The circuit is analyzed by finding its response to a general sinusoidal signal voltage.[6] For example, suppose $v_0(t) = \sin 2\pi ft$. It can be shown that if this voltage wave is applied to the circuit of figure 19.4, the output will be

$$v_1(t) = \frac{1}{\sqrt{1 + (2\pi f)^2 (RC)^2}} \sin(2\pi ft - \theta),$$

where $\tan\theta = 2\pi fRC$. This shows that the magnitude of the signal is sharply attenuated at high frequencies. Thompson understood that the sharp edges of a pulse represented high-frequency components of the signal. Their attenuation explains the

[5]Thompson's analysis actually took account of the propagation of voltage v along the continuous line. He developed the partial differential equation $\partial^2 v/\partial x^2 = RC\partial v/\partial t$ for the voltage at distance x and time t, which is still regarded as the fundamental equation for a line with capacity and resistance. This equation is, as he understood, identical in form to the equation for propagation of heat along an iron rod.

[6]The equations governing the circuit of figure19.4 are $v_0 - v_1 = IR$ and $CI = \mathrm{d}v_1/\mathrm{d}t$, where I is the current. Elimination of I gives

$$\frac{\mathrm{d}v_1}{\mathrm{d}t} = -\frac{1}{RC}(v_1 - v_0).$$

If a constant voltage v_0 is applied at time $t = 0$, then the solution is $v_1(t) = (1 - e^{-t/RC})v_0$.

blurring effect of the transmission, illustrated in figure 19.3. From this blurring effect he estimated that the first transatlantic cable could sustain a transmission rate of 4,000 words per day (that is, about 3 words per minute), which was just enough to render the project economically feasible.

Thompson at least partially discovered and used two of the most important properties of frequency: the invariance property of sinusoids and the Fourier representation. His work was proof that even the early theory had enormous practical power, which could transform the world of communication.

19.3 Fourier Series

A function $x(t)$ is **periodic** with period T if $x(t+T) = x(t)$ for all t, $-\infty < t < \infty$, and if T is the smallest positive value with this property. Two periodic functions are the sine and cosine functions $\sin 2\pi t/T$ and $\cos 2\pi t/T$. These are said to be sinusoidal functions with period T and frequency $f_0 = 1/T$. If, as we almost always suppose, t represents time, frequency has units of cycles per second (or equivalently, units of hertz). One complete cycle of the sinusoidal function is completed every $T = 1/f_0$ seconds.

Frequency is alternatively expressed in terms of radians per second. In these units the frequency associated with a period of duration T is $\omega_0 = 2\pi/T$ radians per second, and the basic sine function with period T is $\sin \omega_0 t$. Radians are often used because the notation is somewhat cleaner without the factors of 2π.

The **harmonics** of the basic sinusoidal functions of period T are the sine and cosine functions with frequencies that are integer multiples of the basic frequency f_0. For example, $\sin 2\pi(2f_0)t = \sin 2\pi(2t/T)$ completes two cycles every T seconds.

The amazing theory of Fourier is that any[7] function of period T can be decomposed into a series of sinusoidal functions of period T and harmonics of these sinusoids. Specifically, the **Fourier series** of a function $x(t)$ with period T uses sines and cosines, and is written as

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi n f_0 t) + \sum_{n=1}^{\infty} b_n \sin(2\pi n f_0 t). \tag{19.1}$$

The coefficients a_0, and a_n, b_n for $n \geq 1$ can be found by simple formulas as well, as shown in exercise 2.

Imaginary Exponentials

A shorthand way of expressing combinations of sines and cosines is to use the special identity

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

where i is the imaginary number $i = \sqrt{-1}$. It follows, likewise, that $e^{-i\theta} = \cos\theta - i\sin\theta$. Complex exponentials obey the rule $e^{a+i\theta} = e^a[\cos\theta + i\sin\theta]$.

[7] The function must satisfy certain technical assumptions concerning boundedness and continuity.

Fourier series can be expressed in terms of complex exponentials as

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{i2\pi n f_0 t}, \tag{19.2}$$

where again $f_0 = 1/T$. In this expression the coefficients c_n may be complex numbers.

This more compact representation of the Fourier series is, of course, equivalent to the earlier one. In fact the coefficients are related by

$$\begin{aligned} a_0 &= 2c_0 \\ a_n &= c_n + c_{-n} \\ b_n &= \frac{c_n - c_{-n}}{i}. \end{aligned}$$

19.4 The Fourier Transform

The Fourier series can be extended to nonperiodic functions by letting the period T tend to infinity. The result is termed the **Fourier transform**, and because of its generality, it is today used more often than Fourier series. Indeed, it is the workhorse of almost all frequency analyses.

The Fourier transform representation of a function $x(t)$ is

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{i2\pi ft} \mathrm{d}f. \tag{19.3}$$

Since $e^{i2\pi ft} = \cos 2\pi ft + i \sin 2\pi ft$, the Fourier transform representation of the function $x(t)$ is in terms of sinusoids. The function $X(f)$ is itself termed the **Fourier transform** or alternatively the **frequency spectrum** of $x(t)$ and is analogous to the coefficients of a Fourier series, for $X(f)$ gives the weight to be applied to the sinusoidal term $e^{i2\pi ft}$. A Fourier transform may have $|X(f)| > 0$ for almost all frequencies, since all frequencies can be regarded as harmonics when the period approaches infinity and the corresponding basic frequency approaches zero.

The Fourier transform can be found explicitly from the equation

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-i2\pi ft} \mathrm{d}t. \tag{19.4}$$

The relationship between the original function and its transform is therefore defined by the symmetric pair

$$\begin{aligned} X(f) &= \int_{-\infty}^{\infty} x(t) e^{-i2\pi ft} \mathrm{d}t \\ x(t) &= \int_{-\infty}^{\infty} X(f) e^{i2\pi ft} \mathrm{d}f. \end{aligned}$$

Tables of such pairs for various functions are available in reference texts.

Example 19.1 (A pulse). Consider the unit pulse $x(t)$ of width T centered at $t = 0$ shown in the top portion of figure 19.5. The Fourier transform of this pulse is

$$
\begin{aligned}
X(f) &= \int_{-T/2}^{T/2} e^{-i2\pi ft}\,\mathrm{d}t \\
&= \int_{T/2}^{T/2} [\cos 2\pi ft - i \sin 2\pi ft]\mathrm{d}t.
\end{aligned} \qquad (19.5)
$$

Since the sine is antisymmetric about $t = 0$, the imaginary part of the integral is zero. The remaining integral is easily found to be

$$X(f) = \frac{\sin(\pi f T)}{\pi f}.$$

This function of f is shown in the bottom portion of figure19.5. This example is useful in many applications. Notice that as the width of the pulse is decreased (T made smaller), the Fourier transform widens. This is a general characteristic of functions and their Fourier transforms: narrower functions produce wider transforms.

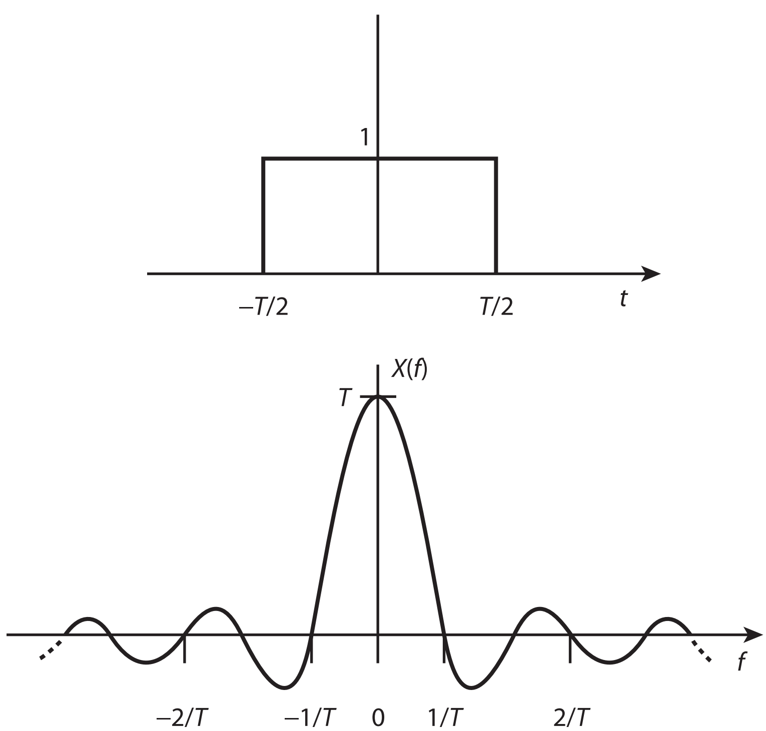


FIGURE 19.5 A square pulse and its transform. A single pulse of finite duration has a Fourier transform that extends over the entire range of frequencies.

Energy Distribution

An important quantity in physical communication systems is the energy represented by a signal, for that energy must be supplied by some physical source. Generally, energy over a short period is proportional to the square of the signal, and hence the total energy required by a signal function $x(t)$ is proportional to the integral

$$E = \int_{-\infty}^{\infty} x(t)^2 \mathrm{d}t.$$

This energy is distributed among frequencies, and that distribution is directly represented by the Fourier transform of the signal. This connection is established by a series of simple steps, manipulating the formula for the total energy.

$$\begin{aligned} E &= \int_{-\infty}^{\infty} x(t)^2 \mathrm{d}t \\ &= \int_{-\infty}^{\infty} x(t) \left[\int_{-\infty}^{\infty} X(f) e^{i2\pi ft} \mathrm{d}f \right] \mathrm{d}t \\ &= \int_{-\infty}^{\infty} X(f) \left[\int_{-\infty}^{\infty} x(t) e^{i2\pi ft} \mathrm{d}t \right] \mathrm{d}f \\ &= \int_{-\infty}^{\infty} X(f) X^*(f) \mathrm{d}f \\ &= \int_{-\infty}^{\infty} |X(f)|^2 \mathrm{d}f. \end{aligned} \tag{19.6}$$

where X^* denotes the complex conjugate of X. This expression for energy is known as **Parseval's theorem** or **Rayleigh's energy theorem**. It shows that the energy distribution among frequencies is given by the function $|X(f)|^2$, which is called the **energy spectral density** of $x(t)$.

For the pulse of example 19.1, the energy spectral density is

$$|X(f)|^2 = \left[\frac{\sin(\pi Tf)}{\pi f} \right]^2,$$

which indicates that the energy is concentrated around $f = 0$ but falls off with higher frequencies.

Since $X(-f) = X^*(f)$, it follows that $|X(-f)|^2 = |X(f)|^2$. Hence, the energy can alternatively be expressed as

$$E = 2 \int_{0}^{\infty} |X(f)|^2 \mathrm{d}f,$$

where the factor 2 is due to writing the integral from 0 to infinity rather than from minus infinity to infinity.

The Fourier transform is a basic tool of frequency analysis. Its power was exhibited by its fundamental role in the development of the telegraph, and it is central in the remaining chapters.

19.5 Thomas Edison and the Telegraph

The discovery of electricity, and especially electromagnetism, spawned a flurry of invention activity beginning in the mid-1800s. The telegraph, being highly visible and commercially successful, attracted young inventive minds like a magnet attracts iron filings.

Among those attracted by the promise and technical excitement of the telegraph was the energetic young man Thomas Alva Edison, who began work as a telegraph operator in 1863 when he was sixteen. He was an unusual operator, however. Frequently he played practical jokes on coworkers (such as wiring a water bucket, from which workers drank, to a high-voltage source), and his energies often drifted away from routine keying and interpretation of messages to the study of telegraph technology. Edison did not stay long in any one location, but rather became a transient operator, traveling from city to city, typically playing jokes on his new colleagues and irritating his supervisors with his inventions. Yet he was an excellent telegrapher, certainly one of the fastest Morse code operators in the country.

One project that fascinated him was the possibility of designing a duplex, which would allow simultaneous transmission of messages in both directions along a single telegraph line. Most of his supervisors dismissed his excursions into the duplex idea as a fruitless waste of time.

Eventually, he gave up telegraph operation to become a full-time inventor. His initial laboratory consisted of space in the shop of Charles Williams, Jr. on Court Street in Boston. Mr. Williams sold electric equipment and often rented laboratory space to young inventors. It was here that Edison first constructed a working duplex using polarized electromagnets (although he was not the first to do so).

Edison's first commercially successful inventions were the quadruplex, improved gold price tickers, and improved stock tickers—all related to the telegraph. As we shall see, his contributions, built on the experience gained while working on the telegraph, played a fundamental role in the advancement of communication.

19.6 Bell and the Telephone

Alexander Graham Bell was born March 3, 1847, in Edinburgh, Scotland. He was exposed to voice training and the synthetic reproduction of voice during his entire youth. Both his father and his grandfather made their livings by the study and teaching of speech and elocution. In 1864 Bell's father, Alexander Melville Bell, developed a universal phonetic alphabet called "Visible Speech," the symbols of which were representations of the configurations of the mouth and tongue that produced the corresponding sounds. This was a major achievement, for several leading phoneticians had failed in their attempts to create such a universal alphabet. It also led to a unique and effective method for teaching deaf people to speak, and Alexander Graham Bell became an excellent instructor in the method.

It was a year later, in 1865, that Alexander Graham Bell, at nineteen, began to experiment with tuning forks. He found that by holding a tuning fork in front of his mouth as he spoke a vowel, the tuning fork of proper pitch would resonate. He soon discovered that each vowel corresponded to a few specific tones. His friend Alexander Ellis told him that these experiments were duplications of earlier experiments by the great German physicist Hermann von Helmholtz, and so Bell studied Helmholtz's book *On the Sensations of Tone*.

Based on his knowledge of Helmholtz's experiments, Bell conceived the idea of sending several messages simultaneously over the telegraph by means of a "harmonic telegraph" in which each of several messages would be sent in Morse code but each using a different pulse tone. To pursue this idea, he began to duplicate Helmholtz's experiments with oscillations derived from tuning forks. He soon replaced the tuning forks with metal reeds, which had the advantage of being smaller and being tunable by adjusting the length of the free part of the reed. A working harmonic telegraph was not easily produced, although Bell tried numerous variations.

In 1874 Bell met Thomas Watson at Williams's electric shop on Court Street, where Thomas Edison had set up his laboratory three years earlier. Watson assisted Bell in his attempts to perfect a workable harmonic telegraph using space in the attic of Williams's shop.

A major breakthrough occurred to Bell on June 2, 1875, while he was engaged in his work on the harmonic telegraph. It was important that the transmitter reeds and receiver reeds be tuned identically. Bell typically held a receiver reed to his ear while with a small instrument he adjusted it to match the tone that was being sent by the transmitter in another room. On that particular day, Watson interrupted the vibration of the transmitter reed to make an adjustment, and Bell, with his ear to the receiver reed, heard sounds associated with Watson's manipulations of the transmitter reed. Bell immediately knew that that was of fundamental significance. He understood then that the wires could transmit, not only tones, but complex sounds. He knew then that a telephone could be built.

What had occurred was the demonstration of the converse of the familiar principle of electromagnets. If a base current flows though the coil, and the armature is forcibly moved (as by impressed sound waves on a reed armature), the current in the coil will respond in accord with the armature movements. This is a manifestation of the dual aspect of electromagnetism: changing current causes armature movement, and armature movement causes changing current. In this context it was a brilliant discovery.

Bell and Watson set about to construct the telephone. The receiver and the transmitter were identical, each consisting of a stretched parchment drumhead with the free end of a harmonic transmitter reed attached at the center of the drumhead. When Watson spoke toward the transmitting drumhead, the sound would be transmitted electrically to another room, where Bell could hear Watson's voice reproduced by the drumhead of the receiver. At least that was the concept. Unfortunately, it did not work well because the feeble current generated by the transmitter was not sufficient to vigorously drive the receiver, so that only a weak and vague response to Watson's voice could be heard. Nevertheless, this primitive telephone launched a transition in understanding. Arbitrary sounds could be sent as electrical signals, simply by electromagnetism.

Bell filed for a patent on his device on February 14, 1876. A diagram in the patent, shown in figure 19.6, makes clear how his first telephone was constructed. The transmitter reed's vibration induces a corresponding vibration in the transmitter coil, which then proceeds to the receiver coil, where it causes the receiver reed to vibrate in sympathy to the original reed. Another inventor, Elisha Gray, filed a caveat (a preliminary patent filing) for a telephone two hours after Bell filed his patent application.

The famous telephone conversation in which Bell entreated, "Mr. Watson. Come here. I want to see you," did not occur until about a month later, on March 10 of the

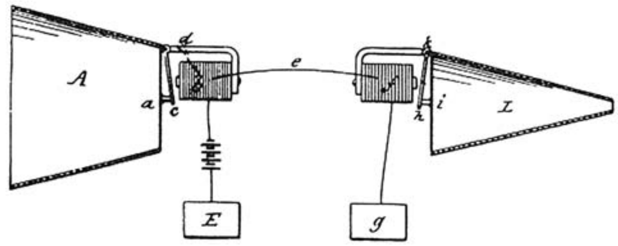


FIGURE 19.6 A diagram from Bell's patent application. Sound waves from a speaker move the transmitting diaphragm and induce a feeble varying current in the coil of the transmitting electromagnet. This current is carried to the receiving electromagnet, which moves the receiver diaphragm.

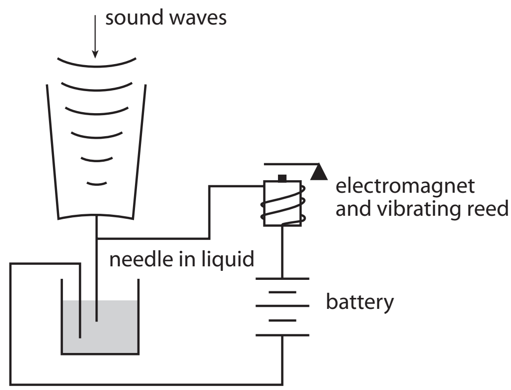
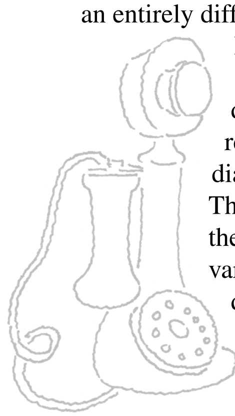


FIGURE 19.7 Bell's variable resistance. The diaphragm moves a needle up and down in the acidic liquid, causing the resistance between the needle and a fixed brass rod to vary according to the sound patterns impressed on the diaphragm.

same year, and it was made over a telephone whose transmitter was constructed on an entirely different principle—that of a microphone.



It was evident that a major problem with the original design was that the current generated by an electromagnet directly connected to a diaphragm was not sufficient to produce intelligible speech at the receiver. The breakthrough modification connected the transmitting diaphragm to a device whose resistance varied as the diaphragm moved. This resistive device was then part of a high-voltage circuit that drove the receiving electromagnet. The setup is shown in figure 19.7. Bell's variable resistance device (or microphone) consisted of a small dish of dilute sulfuric acid into which a brass tube was inserted to serve as one side of the circuit. A needle was connected to the transmitting diaphragm, and the tip of this needle was also inserted into the liquid. As the needle mimicked the vertical vibrations of the diaphragm and slid up and down in the liquid, the needle surface exposed to the liquid varied, and consequently the resistance between the needle and the brass tube was continuously modified. Thus the tiny vibrations of the diagram controlled a large current through an electromagnet that vibrated a reed or other diaphragm. It was with this arrangement that the famous first phone conversation between Bell and Watson took place. Strangely, however, the incident was not reported by either Bell or Watson for more than 10 years.

Controversy and intrigue surrounded the original patent application. The main body of the application says nothing about the use of variable resistance; however, a relatively short paragraph, almost an afterthought, mentions the possibility of using a liquid variable resistance. On the other hand, Elisha Gray's filing fully describes this possibility. Yet there is some evidence that Bell had conceived the concept much earlier. The controversy and numerous other patent disputes, by inventors who apparently had working telephones prior to Bell's, were the subject of much litigation, with one case being finally resolved in the U.S. Supreme Court, which found in favor of Bell in a split decision. A logical inference from these favorable judgments is that Bell was likely very charming and very lucky, as well as hardworking. As an interesting side note, Bell and Gray had earlier each filed patent applications for the harmonic telegraph, with Gray being two days ahead of Bell, and Gray eventually won that patent.

The liquid variable resistance was not commercially practical, and other inventors soon devised superior devices. Indeed it was Thomas Edison who developed the carbon microphone that is still in common use. In the basic design as a diaphragm moves in response to sound waves, it compresses a little space that is filled with small granules of carbon. As the granules are pressed more tightly together, the resistance through them decreases.

19.7 Lessons in Frequency

The saga of Bell's researches, and those of his contemporaries working on similar paths, brought forth an improved mastery of frequency. Two major ideas seem to have emerged. First was the implicit understanding that the shape of a sinusoidal signal is preserved as it passes through electric circuits. Indeed this was the principle underlying the harmonic telegraph.

A second principle—that of a tuned circuit—also was only partly known to the early telephone researchers. These researchers generated sinusoidal signals mechanically, and detected them mechanically as well, with tuning forks or reeds tuned to match those at the transmitter. These practical inventors had not yet discovered that tuning itself could be done with electrical components.

Perhaps the greatest leap in understanding was the recognition that complex sounds, such as those of the human voice, can be thought of as patterns of a single quantity whose amplitude varies; and this pattern can be transcribed as mechanical movements or fluctuations of electrical current, and hence transported over great distances and then converted back to duplicate the original sound. It is perhaps difficult for us to imagine a state of knowledge in which this simple fact is not known. Certainly musicians and physicists knew that sound comes from mechanical movements, but although it was known for a single tone, it was apparently unclear that the complexity of sound could be regarded as a single pattern. This lack of understanding perhaps explains Bell's first conception of a telephone—the harp phone—consisting of a series of reed transmitters with different closely spaced frequencies. A voice projected onto the reeds would stimulate each according to the frequency makeup of the sound, and then be transported to the corresponding array of tuned receiver reeds, where the voice would be reasonably reconstituted as the combination of their separate vibrations. Bell knew that all of these separate tones could be transported over a single wire to the array of tuned receivers, but he did not understand that they did not have to be impressed in their separate identities. However, because of its complexity. Bell never built the harp phone. The original telephone of figure 19.6 is so simple, it is hard to imagine that most of Bell's investigations were attempts to separate and then later gather the individual frequencies.

Shortly after the introduction of the telephone, Thomas Edison created what came to be his favorite invention: the phonograph. In original form, it was not even electric. Edison had been experimenting in his Menlo Park, New Jersey, laboratory with modifications to the telephone, including his carbon microphone, and hence had telephone components readily available. One day in 1877, Edison attached a pin to a diaphragm, and while shouting the word "halloo" onto the diaphragm, he pulled a strip of waxed paper past the pin to record the vibrations as a groove on the paper. He then placed the pin in the groove and pulled the paper through again. He heard a faint "halloo," and the concept of the phonograph was born. Edison was amazed that his voice was captured with that thin wavy scratch. This was one of the first observations of a human speech waveform.

The phonograph was immediately popular, due in part to that little wavy groove that amazed, mystified, and delighted average citizens as much as it did Edison. The breakthrough in understanding could be appreciated and enjoyed by everyone. Indeed, it was upon the introduction of the phonograph that Edison was fondly referred to as the "Wizard of Menlo Park."

In terms of its pure technology, Bell's telephone represents an interesting perplexity. The fundamental observation that electromagnetism works both to generate current and to respond to it enabled his first telephone. That was the essential idea described in his patent, and it clearly was an outstanding achievement. Yet the first practical telephone, constructed a month or so later and capable of producing truly audible sounds at the receiver, did not use that principle; and neither do modern telephones. Instead, a different principle was used—one we find in almost every major technological idea that has advanced the mastery of frequency—the principle of a small signal controlling a larger one (as with Morse's electromagnetic relays). The liquid microphone embodies that idea in Bell's telephone. The tiny vibrations of the transmitting diaphragm are used, by way of the variable resistance, to control the large current in the receiver circuit. It is not certain when Bell conceived of that idea or when he understood its true significance (for he continued to experiment with the earlier idea), but it was that principle that was most profound.

Later, secure with monies generated by the telephone, Bell continued to invent. He invented the hydrofoil ship, the box-shaped airplane wing, and several other important innovations. His favorite, however, which he considered "his most important invention," was what he called the "photophone," by which voice signals were transmitted on a light beam.[8] The device relied on the element selenium, which changes resistance according to the amount of light to which it is exposed. The variable resistance is used to convert the small light signal to a large current in the receiver!

[8]The intensity of the beam was changed by reflecting the beam from a diaphragm controlled by voice. As the diaphragm became more convex, the reflected beam was spread and hence was less intense over any given receiving angle.

19.8 EXERCISES

1. (Alternate form) An alternate version of the Fourier transform uses radians per unit time (ω) rather than cycles per unit time (f). In general, $\omega = 2\pi f$. The Fourier transform of $x(t)$ based on ω is

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t}\mathrm{d}t.$$

   (a) Show that

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{i\omega t}\mathrm{d}\omega.$$

   (b) Show that

$$E = \frac{1}{2\pi}\int_{-\infty}^{\infty} |X(\omega)|^2\mathrm{d}\omega.$$

2. (Fourier coefficients) It is easily seen from figure19.8 that the integral over a period of the product of a sinusoid and its first harmonic is zero. More generally,

$$\int_0^T \sin 2\pi m f_0 t \; \sin 2\pi n f_0 t \; \mathrm{d}t = \begin{cases} 0, & m \neq n \\ T/2 & m = n. \end{cases}$$

$$\int_0^T \cos 2\pi m f_0 t \; \cos 2\pi n f_0 t \; \mathrm{d}t = \begin{cases} 0, & m \neq n \\ T/2 & m = n. \end{cases}$$

$$\int_0^T \sin 2\pi m f_0 t \; \cos 2\pi n f_0 t \; \mathrm{d}t = 0, \quad \text{all } m \text{ and } n.$$

By using these orthogonality relations, show that the coefficients of the Fourier series

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi n f_0 t) + \sum_{n=1}^{\infty} b_n \sin(2\pi n f_0 t)$$

are

$$a_0 = \frac{2}{T}\int_0^T x(t)\mathrm{d}t$$

$$a_n = \frac{2}{T}\int_0^T x(t)\cos(2\pi n f_0 t)\,\mathrm{d}t$$

$$b_n = \frac{2}{T}\int_0^T x(t)\sin(2\pi n f_0 t)\,\mathrm{d}t.$$

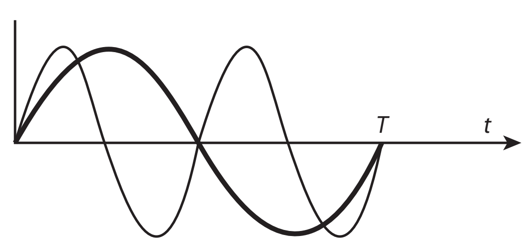


FIGURE 19.8 Orthogonality of a sinusoid and its harmonic.

3. (Square wave series) Find the Fourier series of the square wave with period T defined on $0 \le t \le T$ by

$$x(t) = \begin{cases} 1, & 0 \le t < T/2 \\ -1, & T/2 \le t < T, \end{cases}$$

and extended periodically for all t.

4. (Exponential decay) Find the Fourier transform of

$$x(t) = \begin{cases} 0, & t < 0 \\ e^{-at}, & t \ge 0. \end{cases}$$

5. (Fourier identities) Suppose that the Fourier transforms of $x(t)$ and $y(t)$ are $X(f)$ and $Y(f)$, respectively. Find the Fourier transform of the following signals.
   (a) $ax(t)$
   (b) $x(t) + y(t)$
   (c) $x(at)$
   (d) $x(t + a)$
   (e) $z(t) = \int_{-\infty}^{\infty} x(t - \tau)y(\tau)\mathrm{d}\tau$.

6. (Cosine pulse) Consider the cosine pulse that is zero except in the interval $-T/2 < t < T/2$, where it is $x(t) = \cos 2\pi f_0 t$ with $f_0 = 1/2T$.
   (a) Sketch the pulse shape.
   (b) Write the integral expression for the Fourier transform of $x(t)$.
   (c) Can you infer that you only need to integrate the product of two cosines?
   (d) Find the Fourier transform. Hint: You may find it useful to use the identities

$$2\cos A \cos B = \cos(A + B) + \cos(A - B)$$
$$\sin(A + B) = \sin A \cos B + \cos A \sin B$$
$$\sin(A - B) = \sin A \cos B - \cos A \sin B.$$

7. (Pulse energy) Using Rayleigh's energy theorem applied to the pulse of example 19.1 show that

$$\int_0^{\infty} \frac{\sin^2 x}{x^2}\mathrm{d}x = \frac{\pi}{2}.$$

8. (More Fourier identities) If $X(f)$ is the Fourier transform of $x(t)$, find the transform of
   (a) $\dfrac{\mathrm{d}x(t)}{\mathrm{d}t}$
   (b) $tx(t)$.

9. (An integral*) Evaluate the integral

$$\int_0^{\infty} \frac{\cos\omega}{1 + \omega^2}\mathrm{d}\omega$$

by considering the Fourier transform of

$$x(t) = \begin{cases} 0, & t < 0 \\ e^{-t}, & 0 \le t \le 1. \end{cases}$$

Hint: Use exercise 1 and first consider $y(t) = e^{-t}$ for $t \ge 0$, $y(t) = 0$ for $t < 0$.

19.9 Bibliography

The colorful history of the early telegraph is presented in [1], [2], [3], [4]. An excellent popular textbook treatment of the Fourier transform, including a brief biography of Joseph Fourier, is [5]. A general history of early electrical discovery and engineering is [6]. Histories of Bell's telephone work are [7] and [8]. See also [9] for an introduction to telephone technology and its history. A detailed study of the controversy surrounding the Bell telephone patent is contained in [10]. There are several biographies of Edison; two that were especially useful in preparation of this chapter are [11] and [12].

References

[1] Standage, Tom. *The Victorian Internet*. New York: Berkley Books, 1998.
[2] Wilson, Geoffrey. *The Old Telegraphs*. London: Phillimore, 1976.
[3] Clarke, Arthur. *Voice across the Sea*. London: William Luscombe, 1958.
[4] Coe, Lewis. *The Telegraph: A History of Morse's Invention and Its Predecessors in the United States*. Jefferson, N.C.: McFarland, 1993.
[5] Bracewell, Ronald N. *The Fourier Transform and Its Applications*. 3rd ed. New York: McGraw-Hill, 1986.
[6] Skilling, H. H. *Exploring Electricity: Man's Unfinished Quest*. New York: Ronald Press, 1948.
[7] Bruce, Robert V. *Bell: Alexander Graham Bell and the Conquest of Solitude*. Boston: Little, Brown, 1973.
[8] Frailey, Jarrell D., and James M. Velayas. *In the Spirit of Service: Telecommunications from the Founders to the Future*. St. Louis: Columbia Creek Publishing, 1993.
[9] Noll, A. Michael. *Introduction to Telephones & Telephone Systems*. 2nd ed. Norwood, Mass.: Artech House, 1991.
[10] Aitken, William. *Who Invented the Telephone?* London: Blackie and Sons, 1939.
[11] Adair, Gene. *Thomas Alva Edison: Inventing the Electric Age*. New York: Oxford University Press, 1996.
[12] Israel, Paul. *Edison: A Life of Invention*. New York: Wiley, 1998.