Contents

Chapter 1

Radio Mathematics, Oscillators, and Transmitters

> What is the soul of mathematics, and to what wavelength must our souls be tuned to catch its message?
>
> —David Eugene Smith (1860–1944), speaking in 1921 as the retiring president of the American Mathematical Society. Just a few years earlier, this metaphor would have been meaningless to almost everybody.[1]

1.1 Kirchhoff's Laws and FitzGerald's Oscillating Circuit

As you start reading this first chapter (or at any time as you read this book), take a parallel look at the appendix. That will give you an appreciation for the central role high-frequency[2] sinusoidal oscillations play in radio, starting at the transmitter. (Oscillators are in receivers, too, as you'll see in subsequent chapters.) It was understood, right from the moment Maxwell published his *Treatise* (when all that he had written was still pretty much *theory*) that the crucial next step to elevate speculative theory to hard fact was to actually generate the oscillating electromagnetic waves the field equations predict. How to do that?

The key idea for the first (and eventually successful) approach to generating radio frequency (rf) waves came in 1883 from the Irish

1 Smith's address is reprinted under the title "Religio Mathematici" in the October 1921 issue of the *American Mathematical Monthly*.

2 "High-frequency" is dictated by the quarter-wavelength $\left(\frac{1}{4}\lambda\right)$ requirement (discussed in the appendix) for the transmitter antenna. To make $\frac{1}{4}\lambda$ a physically reasonable value, the frequency has to be "high." For example, to build a radio antenna transmitting at the power-line frequency of 60 Hz would be ridiculous, as the wavelength at that frequency is 5 million meters. A $\frac{1}{4}\lambda$ antenna at 60 Hz would be 777 miles high, more than three times the orbital height of the International Space Station!

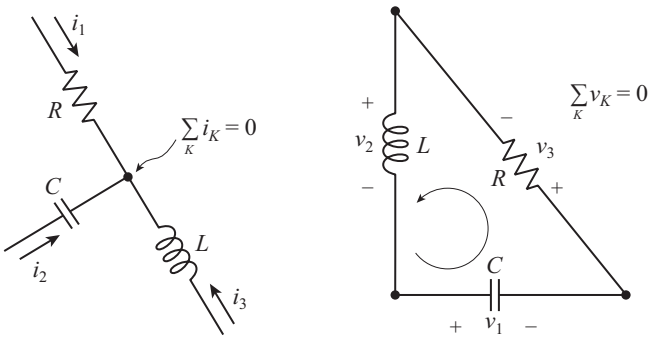

FIGURE 1.1.1. Kirchhoff's two circuit laws.

physicist George Francis FitzGerald (1851–1901). FitzGerald suggested charging a capacitor (with the aid of a static electricity generator) to a high voltage and then letting it discharge through an inductive circuit. (This is actually only slightly more complicated than the circuits of Professors Twombly and Tweedle in the preface.) FitzGerald suggested that oscillations with a wavelength of 10 m (meters) might be achieved (a frequency of 30 MHz (megahertz)—30 million cycles per second. To understand what FitzGerald was talking about requires us first to establish the two fundamental laws obeyed by the electrical circuits you'll find in all radio electronics. These are *Kirchhoff's laws*—after the German physicist Gustav Robert Kirchhoff (1824–1887) who formulated them in 1845—which are simply the laws of conservation of energy and the conservation of electric charge. With reference to Figure 1.1.1, we have

Kirchhoff's current law: The sum of the currents into any node (a point where components are connected together) is zero. This is conservation of electric charge. In other words, charge transported into any node by a current is transported out of the node by another current.

Kirchhoff's voltage law: The sum of the voltage drops (or of the voltage increases) around any closed-loop path in a circuit is zero. This is conservation of energy. You can see this by recalling that voltage is energy per unit charge, and a voltage drop is the energy required to transport a unit charge through the electric field that exists in a component, and so the law says the total energy to go around a closed

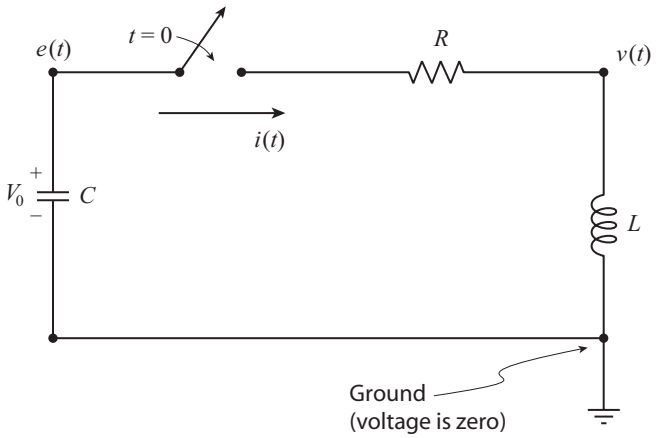

FIGURE 1.1.2. FitzGerald's oscillating circuit.

loop is zero. If it were *not* zero, then we could endlessly transport charge around a closed loop *in the sense for which the energy required is negative* and so become rich selling the gained energy to the local power company! (You'll believe *that* only if you believe in the possibility of a perpetual motion machine.)

Now we can understand what FitzGerald was suggesting. Figure 1.1.2 shows his circuit, with the capacitor C charged to V_0 volts. At $t=0$ we close the switch, and so now there is a path through the resistor R and the inductor L in which the current $i(t)$ can flow.

Just before we close the switch, the stored energy in the circuit is just the energy in the electric field of C ($i(t)=0$ for $t<0$, and so, as Professor Tweedle states at the end of the preface, *initially* there is no stored energy in the magnetic field of L, because there *is* no magnetic field in L for $t<0$).

If $W(t)$ is the total energy in the circuit, then in general we have

$$W(t)=\frac{1}{2}Ce^2(t)+\frac{1}{2}Li^2(t),$$

and if we differentiate with respect to t,

$$\frac{dW}{dt}=Ce\frac{de}{dt}+Li\frac{di}{dt}.$$

But since

$$i(t) = -C\frac{de}{dt},$$

then

$$\frac{dW}{dt} = Ce\left(-\frac{i}{C}\right) + Li\frac{di}{dt} = -i\left(e - L\frac{di}{dt}\right).$$

Now, as the inductor voltage $v(t)$ is

$$v(t) = L\frac{di}{dt},$$

we have

$$\frac{dW}{dt} = -i(e - v).$$

Since Ohm's law says

$$i(t) = \frac{e - v}{R},$$

we have $(e - v) = iR$, and therefore,

$$\frac{dW}{dt} = -i^2R < 0,$$

because no matter what $i(t)$ is, $i^2 \geq 0$. Thus, $\frac{dW}{dt}$ is always negative if we assume $R \geq 0$. This may appear to be a trivial assumption, as *of course* R is positive, right? After all, just go into a store selling electrical parts and ask for a box of negative resistors, and see what the clerk says! *But*, in fact, as we'll get to soon when we discuss how Hardy's friend Littlewood tackled the Van der Pol equation, there *is* such a thing as negative resistance, and, in fact, the entire development of modern electronics is based on that fact.

For now, however, in FitzGerald's preelectronic 1883 circuit R has a positive value, and so we see the initial stored energy in the C continuously decreases once the switch is closed. The central issue raised

by FitzGerald, however, was *not* that the energy decreases but rather *how* that decrease occurs. To answer that question, let's look in more detail at $i(t)$. Starting at the ground node (which, by definition is at a voltage of zero) in Figure 1.1.2, let's write Kirchhoff's voltage law as we go around the loop in a clockwise sense (the sum of the voltage drops[3] is zero):

$$-e(t)+i(t)R+L\frac{di}{dt}=0.$$

Differentiating with respect to time,

$$-\frac{de}{dt}+R\frac{di}{dt}+L\frac{d^2i}{dt^2}=0,$$

or, as we observed before, since

$$\frac{de}{dt}=-\frac{i}{C},$$

we have

$$\frac{i}{C}+R\frac{di}{dt}+L\frac{d^2i}{dt^2}=0,$$

or

$$\frac{d^2i}{dt^2}+\frac{R}{L}\frac{di}{dt}+\frac{1}{LC}i=0.$$

The standard method for solving this second-order differential equation is to assume the solution

$$i(t)=Ie^{st},$$

where s is some constant to be determined. Substituting this assumption back into the differential equation, we get

3 As we travel through the C we experience a voltage *rise* from zero to e, which explains why we write $-e$ as the voltage *drop*.

$$Is^2e^{st} + \frac{R}{L}Ise^{st} + \frac{1}{LC}Ie^{st} = 0,$$

and so, making the obvious cancellations (which explains why this method works!), we get

$$s^2 + \frac{R}{L}s + \frac{1}{LC} = 0,$$

a result that lets us solve for what the constant s actually is (in fact, there are *two* such values):

$$s = \frac{1}{2}\left\{-\frac{R}{L} \pm \sqrt{\left(\frac{R}{L}\right)^2 - \frac{4}{LC}}\right\}.$$

Now, notice that for given values of L and C, if we have R sufficiently small so that

$$\left(\frac{R}{L}\right)^2 < \frac{4}{LC},$$

then with $j = \sqrt{-1}$ we have[4]

$$s = \frac{1}{2}\left\{-\frac{R}{L} \pm j\sqrt{\frac{4}{LC} - \left(\frac{R}{L}\right)^2}\right\},$$

or, more compactly, our two values of s are

$$s_1 = -\frac{R}{2L} + j\omega_0, \quad s_2 = -\frac{R}{2L} - j\omega_0, \quad \omega_0 = \sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2}.$$

Thus, the most general solution for $i(t)$ is

4 Mathematicians almost always write $i = \sqrt{-1}$ and like to joke that electrical engineers write $j = \sqrt{-1}$ because otherwise they'll confuse $\sqrt{-1}$ with electrical currents in their circuits (which are usually written with the symbol i). This, of course, is nonsense of a near-libelous nature—but, I have to admit, it *is* less confusing not to use i for both concepts. So, if mathematicians will let me write $j = \sqrt{-1}$ and reserve i for currents, I will, in turn, promise not to tell any silly mathematician jokes in this book.

$$i(t) = I_1 e^{s_1 t} + I_2 e^{s_2 t}.$$

To calculate what I_1 and I_2 are, we start with the following important fact about inductors: the current in an inductor cannot change instantly, which follows immediately from its mathematical description. That is, if the current i in inductor could change instantly, then the voltage drop across the inductor would be infinite (because $\frac{di}{dt}$ would be infinite). Engineers and physicists reject the possibility of a physical infinity as nonsense, and so the current in FitzGerald's circuit at $t = 0+$ (immediately after the switch closes) must equal the current at $t = 0-$ (immediately before the switch closes).[5] Since $i(0-) = 0$, then $i(0+) = 0$, too, and we have

$$i(0+) = 0 = I_1 + I_2,$$

and so $I_1 = -I_2 = I$, which gives us

$$i(t) = I(e^{s_1 t} - e^{s_2 t}).$$

To determine what I is, we again use the fact that $i(0+) = 0$, which means (because of Ohm's law) that the voltage drop across R is zero. *That* means, because of Kirchhoff's voltage law, that the initial capacitor voltage V_0 appears across L at $t = 0+$, and so

$$V_0 = L\frac{di}{dt}\Big|_{t=0+},$$

or

$$\frac{di}{dt}\Big|_{t=0+} = \frac{V_0}{L} = I(s_1 e^{s_1 t} - s_2 e^{s_2 t})\Big|_{t=0+} = I(s_1 - s_2),$$

or

$$I = \frac{\frac{V_0}{L}}{s_1 - s_2} = \frac{\frac{V_0}{L}}{j2\omega_0}.$$

5 According to the same sort of argument, it follows that the voltage drop across a capacitor cannot change instantly, as that would require an infinite current.

That is,

$$i(t)=\frac{V_0}{j2\omega_0 L}\left[e^{\left(-\frac{R}{2L}+j\omega_0\right)t}-e^{\left(-\frac{R}{2L}-j\omega_0\right)t}\right]=\frac{V_0e^{-\frac{R}{2L}t}}{j2\omega_0 L}[e^{j\omega_0 t}-e^{-j\omega_0 t}]$$
$$=\frac{V_0e^{-\frac{R}{2L}t}}{j2\omega_0 L}\,j\,2\sin(\omega_0 t),$$

where I've used *Euler's identity*.[6] Thus,

$$i(t)=\frac{V_0e^{-\frac{R}{2L}t}}{\omega_0 L}\sin(\omega_0 t),\ \omega_0=\sqrt{\frac{1}{LC}-\left(\frac{R}{2L}\right)^2},$$

assuming that $R<2\sqrt{\frac{L}{C}}$. The current $i(t)$ is said to be an *alternating current*, popularly known as "ac."

So, FitzGerald was correct in saying his circuit will, if R is sufficiently small, oscillate sinusoidally at a particular frequency determined by the values of the circuit components.[7] Does that, however, mean the circuit will generate rf waves? We can explore that question by re-doing the calculation done by Professors Tweedle and Twombly in the Preface: the evaluation of the heat energy integral

$$\int_0^\infty i^2R\,dt=\left(\frac{V_0}{\omega_0 L}\right)^2 R\int_0^\infty e^{-\frac{R}{L}t}\sin^2(\omega_0 t)\,dt.$$

This is a straightforward (if *slightly* messy) freshman calculus calculation, and I'll let you confirm that its value is $\frac{1}{2}CV_0^2$, precisely the

6 Euler's identity, $e^{jx}=\cos(x)+j\sin(x)$, (due to the Swiss-born mathematician Leonhard Euler (1707–1783), is at the very heart of AM, FM, and SSB radio theory, and we will use it repeatedly in this book.

7 The oscillations are a manifestation of the circuit's stored energy sloshing back and forth between the electric field of the C and the magnetic field of the L. Electrical engineers demonstrate the poetic nature of their souls by picturesquely calling the LC combination a *tank circuit*, a reference to the sloshing of water waves back and forth in a disturbed water tank.

value of the initial stored energy in the capacitor. FitzGerald's circuit, therefore, *as it stands*, is no better than Twombly's in generating radio waves. But, unlike Twombly's, all FitzGerald's circuit needs is one final touch—the addition of an antenna! (This is where you really need to read the appendix, particularly the end of it.)

From the oscillating current in FitzGerald's circuit, the resulting oscillating magnetic field of the L can be coupled via Faraday's electromagnetic induction (as shown in Figure 1.1.3) into the antenna, to serve as the oscillating voltage that drives the conduction electrons in the antenna back and forth. That motion, as explained in the appendix, creates kinks in the electric field in the space around the antenna, kinks which in turn give rise to a Poynting energy-flow vector always directed away from the antenna.

As it stands in Figure 1.1.3, FitzGerald's circuit won't transmit for long, because the initial energy in the C is quickly dissipated as heat in the R and as rf waves from the antenna. The early radio experimenters attempted to keep the oscillations going by periodically injecting new energy into the circuit, by incorporating a repeatedly operating spark gap, reaching speeds of up to 20,000 sparks per second. With each new spark a pulse of energy was injected, and such radio transmitters sounded like machine guns! This was okay for Morse code wireless telegraphy but totally inadequate for use in what would become modern voice-and-music radio, and I'll not pursue that approach to radio in this book.[8]

A much different approach was to introduce a *negative resistance* into the oscillator circuit, to counter the energy loss caused by the positive R and the rf radiation. This was achieved, most importantly, with the invention in 1906–1907 of the triode electronic vacuum tube that so captured Einstein's imagination, but it was preceded in the nineteenth century by the *electric arc*. We'll briefly discuss the arc once we have established more mathematical results in the next section.

8 You can find a detailed mathematical discussion of spark-gap radio in my book *The Science of Radio*, Springer 2001. Such radios are now mostly of historical interest, as spark-gap radio has been illegal since 1923, for reasons based on the mathematics (which itself remains quite interesting).

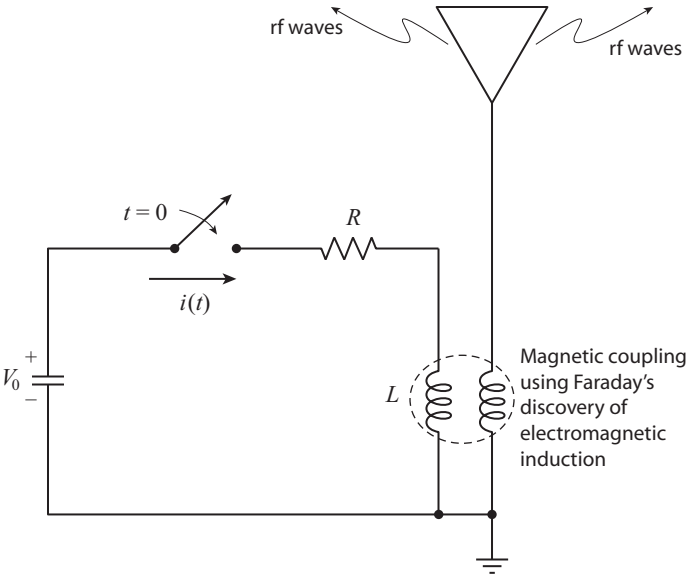

FIGURE 1.1.3. FitzGerald's circuit as a transmitter (the oscillating current $i(t)$ in L creates an oscillating magnetic field that, in turn, creates an oscillating voltage in the antenna which drives conduction electrons in the antenna back-and-forth, creating electric field kinks.

1.2 Laplace Transforms, AC Impedance, and Transfer Functions

We will be concerned in all our discussions of radio with electrical signals that vary *sinusoidally* with time, that is, with signals like $\cos(\omega_0 t)$ and $\sin(\omega_0 t)$, where ω_0 is the *angular frequency* (in radians per second). (As used here, ω_0 is an *arbitrary* frequency and is not the particular ω_0 of the previous section.) If f_0 is the frequency in hertz (what used to be called cycles per second), then $\omega_0 = 2\pi f_0$. AM radio frequencies are in the interval 540 to 1600 kHz (kilohertz),[9] while FM radio operates in the interval 88 to 108 MHz (megahertz). We will find that the differential equations that describe how numerous radio circuits work are *linear*, which means that the sum of two solu-

9 The first radio program Hardy heard was almost certainly broadcast by the BBC London-based station 2LO, which began operating in 1922 at 842 kHz (that is, at 842 kilocycles).

tions to the differential equations is also a solution. Thus, rather than studying the behavior of a circuit in response to, say, a voltage signal like $\cos(\omega_0 t)$ or $\sin(\omega_0 t)$, we can do both problems simultaneously by studying the solution to the complex voltage signal $e^{j\omega_0 t}$, because Euler's identity says $e^{j\omega_0 t} = \cos(\omega_0 t) + j\ \sin(\omega_0 t)$.

This is because the solution for the signal $e^{j\omega_0 t}$ is the sum of the solution to the signal $\cos(\omega_0 t)$ and the solution to the signal $j\sin(\omega_0 t)$. The solution to the signal $j\sin(\omega_0 t)$ will be the solution to $\sin(\omega_0 t)$ multiplied by the constant j (again, by linearity), and so the solution to the signal $\cos(\omega_0 t)$ will be the real part of the solution to $e^{j\omega_0 t}$, and the solution to the signal $\sin(\omega_0 t)$ will be the imaginary part of the solution to $e^{j\omega_0 t}$. This simple idea leads to the enormously useful concept of *ac impedance*, which allows us (for *sinusoidal* time functions) to treat capacitors and inductors as obeying Ohm's law, which up to now has been limited to resistors.

Since $e^{j\omega_0 t} = \cos(\omega_0 t) + j\ sin(\omega_0 t)$, it follows that $e^{-j\omega_0 t} = \cos(-\omega_0 t) + j\ \sin(-\omega_0 t) = \cos(\omega_0 t) - j\ \sin(\omega_0 t)$. Thus, $\cos(\omega_0 t) = \frac{1}{2}[e^{j\omega_0 t} + e^{-j\omega_0 t}]$, and $\sin(\omega_0 t) = \frac{1}{j2}[e^{j\omega_0 t} - e^{-j\omega_0 t}]$, and both of these expressions have simple physical interpretations. In the complex plane, $e^{j\omega_0 t}$ and $e^{-j\omega_0 t}$ are vectors of unit length (because $\cos^2(\omega_0 t) + \sin^2(\omega_0 t) = 1$), making angles $\omega_0 t$ and $-\omega_0 t$ with the real axis, respectively, as shown in Figure 1.2.1a. Indeed, since these two angles increase as t (time) increases, $e^{j\omega_0 t}$ and $e^{-j\omega_0 t}$ are counterrotating vectors, both with real part $\cos(\omega_0 t)$ and with imaginary parts $\sin(\omega_0 t)$ and $-\sin(\omega_0 t)$, respectively. If we sum these two vectors as they rotate, it is obvious their imaginary parts cancel and their real parts add, to give an oscillating result that always lies along the real axis. If, however, we subtract $e^{-j\omega_0 t}$ from $e^{j\omega_0 t}$, we simply multiply $e^{-j\omega_0 t}$ by –1 (which reflects $e^{-j\omega_0 t}$ through the origin) and add, as shown in Figure 1.2.1b. This addition results in the real parts cancelling and the imaginary parts adding, to give us an oscillating result that always lies along the imaginary axis.

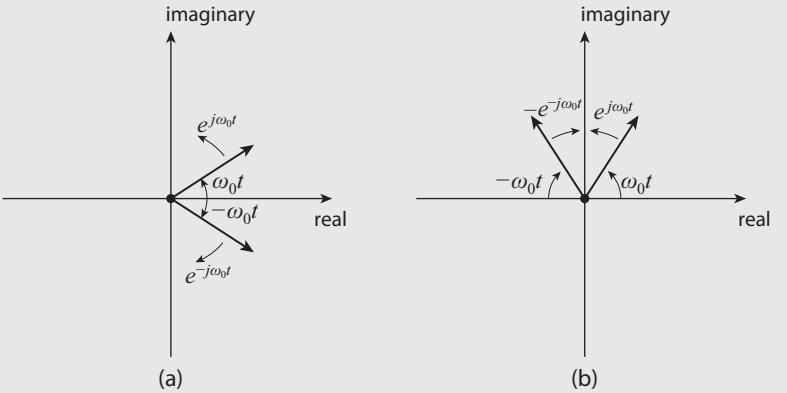

FIGURE 1.2.1. Euler's identity and counterrotating vectors.

To start our development of the impedance concept, consider Fitz-Gerald's series circuit again, but now powered by a complex-valued voltage source, as shown in Figure 1.2.2.

We assume there is, most generally, an initial current $i(0+)$ in the circuit, as well as an initial charge $q(0+)$ in the capacitor. Thus, with $i(t)$ the current for $t \geq 0$, the differential equation that describes the circuit is, using Kirchhoff's voltage loop law (starting at the negative terminal of the voltage source and going clockwise around the loop),

$$-v(t) + iR + \frac{1}{C}\left[\int_0^t i(x)dx + q(0+)\right] + L\frac{di}{dt} = 0,$$

or

$$Ee^{j\omega_0 t} = iR + \frac{1}{C}\left[\int_0^t i(x)dx + q(0+)\right] + L\frac{di}{dt}.$$

Taking the Laplace transform (see the following box) with $a = -j\omega_0$ for the term on the left, we have[10]

10 Notice that on the right we have $\mathcal{L}\{q(0+)\} = \int_0^\infty q(0+)e^{-st}dt = -\frac{q(0+)}{s}e^{-st}\Big|_0^\infty = \frac{q(0+)}{s}$, where, to evaluate the upper limit, $\lim_{s\to\infty} e^{-st} = 0$, as s is defined to have a *positive* real part.

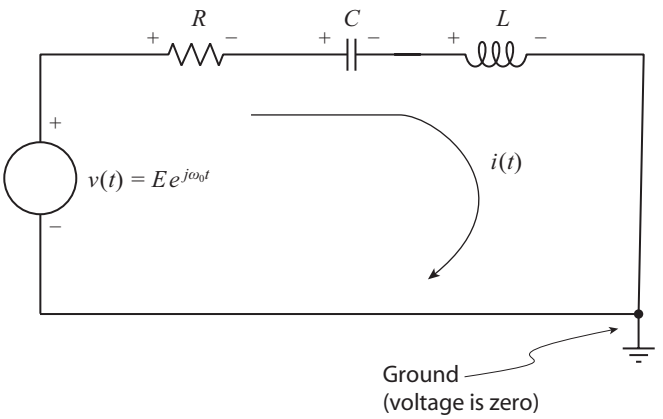

FIGURE 1.2.2. FitzGerald's circuit with a complex-valued voltage source.

$$\frac{E}{s-j\omega_0}=I(s)R+\frac{1}{C}\left[\frac{I(s)}{s}+\frac{q(0+)}{s}\right]+L\left[sI(s)-i(0+)\right]$$

and it is an easy matter to solve for $I(s)$:

$$I(s)=\frac{1}{sL+R+\dfrac{1}{sC}}\left[\frac{E}{s-j\omega_0}+Li(0+)-\frac{q(0+)}{sC}\right].$$

In radio theory we will be interested in time functions $f(t)$ that vanish for $t<0$. Physically, we interpret the instant $t=0$ as when we "turn $f(t)$ on." The Laplace transform of $f(t)$ is $\mathcal{L}\{f(t)\}=F(s)=\int_0^\infty f(t)e^{-st}dt$; the variable s is a complex variable with a positive real part to ensure convergence of the integral, but we can often formally work with the Laplace integral as if s is real. The value of the transform lies in its conversion of certain "complicated" operations in the time domain (like differentiation and definite integration) into "simple" algebraic ones. Specifically, $\mathcal{L}\left\{\dfrac{d}{dt}f(t)\right\}=sF(s)-f(0+)$, and $\mathcal{L}\left\{\int_0^t f(x)dx=\dfrac{1}{s}F(s)\right\}$. Tables of transforms have been created over the decades for

a vast number of time functions, but the one most useful in radio analyses is $\mathcal{L}\{e^{-at}\}=\dfrac{1}{s+a}$, where a is a constant. When $a=0$, this says $\mathcal{L}\{1\}=\dfrac{1}{s}$. More precisely, in the $a=0$ case we are dealing with the function $f(t)=\begin{cases}1, & t>0\\ 0, & t<0\end{cases}$, which is called the *Heaviside step function*, often written as $H(t)$, in honor of Oliver Heaviside (see the appendix), who made extensive use of it. Of course, there are also other important *transform pairs* of $f(t)\leftrightarrow F(s)$, but the exponential time function transform will do 95% of the work for us here.

The factor

$$\frac{1}{sL+R+\dfrac{1}{sC}}=\frac{sC}{s^2LC+RCs+1}=\frac{s}{s^2L+Rs+\dfrac{1}{C}}=\frac{s}{L\left(s^2+\dfrac{R}{L}s+\dfrac{1}{LC}\right)}$$

can be written in the form

$$\frac{s}{L(s-s_1)(s-s_2)},$$

where s_1 and s_2 are each a function of R, L, and C.[11] Thus, the Laplace transform of the current is

$$\begin{aligned}I(s)&=\frac{s}{L(s-s_1)(s-s_2)}\left[\frac{E}{s-j\omega_0}+Li(0+)-\frac{q(0+)}{sC}\right]\\&=\frac{\dfrac{E}{L}s}{(s-s_1)(s-s_2)(s-j\omega_0)}+\frac{si(0+)}{(s-s_1)(s-s_2)}-\frac{q(0+)/LC}{(s-s_1)(s-s_2)}\\&=\frac{\dfrac{E}{L}s+(s-j\omega_0)si(0+)-(s-j\omega_0)\dfrac{q(0+)}{LC}}{(s-s_1)(s-s_2)(s-j\omega_0)},\end{aligned}$$

11 I'll leave it for you to confirm (it's easy!) that s_1 and s_2 are both either *real and negative* or both complex with *negative real parts* for any choice of positive values for R, L, and C. That's all we'll need to know about s_1 and s_2, as you'll soon see.

or

$$I(s)=\frac{\frac{E}{L}s+(s-j\omega_0)\left[si(0+)-\frac{q(0+)}{LC}\right]}{(s-s_1)(s-s_2)(s-j\omega_0)}.$$

If you examine $I(s)$, you see it has the form of a fraction with a numerator that is quadratic in s divided by a denominator that is cubic in s. It is therefore clear that we can write $I(s)$ as the partial-fraction expansion

$$I(s)=\frac{N_1}{s-s_1}+\frac{N_2}{s-s_2}+\frac{N_3}{s-j\omega_0},$$

where N_1, N_2, and N_3 are constants. If we now return to the time domain (using the exponential transform pair), this says

$$i(t)=N_1e^{s_1t}+N_2e^{s_2t}+N_3e^{j\omega_0t}.$$

Since s_1 and s_2 are either both negative or are both complex with negative real parts (see note 11), we see that the first two terms go to zero as $t\to\infty$. These two terms, which disappear with increasing time, represent *transient currents*. The third term, however, does *not* vanish as $t\to\infty$ but endlessly oscillates (because of Euler's identity). This persistent term is called a *steady-state current*. We can calculate N_3 by multiplying through $I(s)$ by the factor $s-j\omega_0$ and then taking the limit $s\to j\omega_0$. That is,

$$N_3=\lim_{s\to j\omega_0}(s-j\omega_0)I(s)=\lim_{s\to j\omega_0}\frac{\frac{E}{L}s}{(s-s_1)(s-s_2)}.$$

Since *by definition*

$$(s-s_1)(s-s_2)=s^2+\frac{R}{L}s+\frac{1}{LC},$$

then

$$N_3=\frac{\frac{E}{L}j\omega_0}{(j\omega_0)^2+\frac{R}{L}j\omega_0+\frac{1}{LC}}=\frac{\frac{E}{L}}{j\omega_0+\frac{R}{L}+\frac{1}{j\omega_0LC}}=\frac{E}{j\omega_0L+R+\frac{1}{j\omega_0C}}.$$

So, the steady-state current (the current after all transients have become insignificant) is

$$i(t)=\frac{Ee^{j\omega_0 t}}{R+j\omega_0 L+\dfrac{1}{j\omega_0 C}},$$

and this current is in response to the voltage $v(t)=Ee^{j\omega_0 t}$. Thus, *for the special case of sinusoids*, we see that we have a result that "looks like" Ohm's law; that is, if we write

$$Z(j\omega_0)=R+j\omega_0 L+\frac{1}{j\omega_0 C}=R+j\left(\omega_0 L-\frac{1}{\omega_0 C}\right)$$

as a sort of "resistance" (radio engineers call the frequency-dependent $Z(j\omega_0)$ the *ac impedance* at frequency ω_0), then for the steady state we have (where the symbols for voltage V and current I are written in uppercase to emphasize we are considering *only sinusoidal* time functions)

$$V(j\omega_0)=Z(j\omega_0)I(j\omega_0).$$

The unit of impedance is ohms, but unlike a resistance, which is purely real, an impedance is generally complex (the imaginary part of Z is called the *reactance*). This result, you'll notice, holds for *any* $i(0+)$ and *any* $q(0+)$; that is, while the initial conditions affect the transient terms, *they play no role in the steady-state term*. If $\omega_0=\dfrac{1}{\sqrt{LC}}$, then $|Z|$ is minimized (equal to R, with zero reactance), and ω_0 is called the *resonant frequency*.

We have the further observation that, at *any* frequency, the ac impedance of a resistor is R, the ac impedance of L at frequency ω is $j\omega L$, and the ac impedance of C at frequency ω is $\dfrac{1}{j\omega C}$. (Notice that I'm now writing ω, not ω_0, since the frequency of the input $v(t)$ is arbitrary, and a subscript is not necessary.) When working with ac impedances we can treat inductors and capacitors, *mathematically*, just like we treat resistors. So, when impedances are in series (as

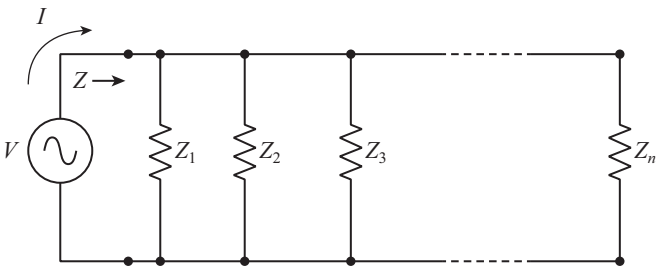

FIGURE 1.2.3. Z is the equivalent impedance of n parallel impedances.

they are in FitzGerald's circuit) they add. When impedances are in parallel their *reciprocals* add, a slightly nonobvious result we can see as follows with reference to Figure 1.2.3.

We have the impedance "seen" by the voltage source V as

$$Z = \frac{V}{I},$$

while from Kirchhoff's current law we have

$$I = \frac{V}{Z_1} + \frac{V}{Z_2} + \frac{V}{Z_3} + \ldots + \frac{V}{Z_n},$$

and so

$$\frac{I}{V} = \frac{1}{Z} = \frac{1}{Z_1} + \frac{1}{Z_2} + \frac{1}{Z_3} + \ldots + \frac{1}{Z_n}.$$

The special case of $n = 2$ leads to the very useful rule that two impedances in parallel are equivalent to their product divided by their sum. When analyzing radio circuits it is helpful to notice that the impedance of a capacitor is very large at low frequencies (infinite at zero frequency, or direct current (dc)) but tends to zero as the frequency increases, while the opposite is true for an inductor. (In the next section we'll use the fact that at $\omega = 0$ the dc resistance of an ideal inductor is zero, while the ac impedance can be quite large for any high-frequency energy that may also be present.)

The frequency behaviors of inductors and capacitors can be used to construct circuits that are of central importance in radio. As an example, consider the circuit of Figure 1.2.4. To emphasize that we

are assuming sinusoidal voltages and currents *only*, I've written the input and output voltages in uppercase letters showing explicit dependence on the frequency variable ω (and not in lowercase as arbitrary functions of the time variable, t). You'll see this circuit again, later in this chapter, where I'll show you how it can be used to build an oscillator. For now, to support that discussion we'll need to know what electrical engineers call the *transfer function* $H(j\omega)$ of the circuit; that is, we'll now calculate

$$H(j\omega)=\frac{V_o(j\omega)}{V_i(j\omega)}.$$

A systematic way of calculating $H(j\omega)$ is based on the clever idea of *loop currents*, labeled as $I_1(j\omega)$ and $I_2(j\omega)$ in Figure 1.2.4. The loop-current approach to writing Kirchhoff's voltage loop law was introduced into circuit theory by Maxwell in his 1873 *Treatise*, and it is now a routine part of electrical engineering. The loop currents, *individually*, are fictitious, but they combine to give the actual currents in each component. For example, the current in the left C is I_1-I_2 downward (or I_2-I_1 upward), while the current in the right C is I_2 (to the right). The physical significance of I_1 is that it's the current that must be supplied from whatever is the source of the input voltage V_i. Writing Kirchhoff's voltage loop equations for the two loops in Figure 1.2.4, we have

$$-V_i+I_1R+\frac{1}{j\omega C}(I_1-I_2)=0,$$

and

$$\frac{1}{j\omega C}I_2+I_2R+\frac{1}{j\omega C}(I_2-I_1)=0,$$

which can be written in the form demanded by *Cramer's rule*[12] for solving these two simultaneous algebraic equations for I_1 and I_2:

12 After the Swiss mathematician Gabriel Cramer (1704–1752). Cramer published the rule in 1750, but in fact it had appeared two years earlier in a posthumously published work by the Scottish mathematician Colin MacLaurin (1698–1746).

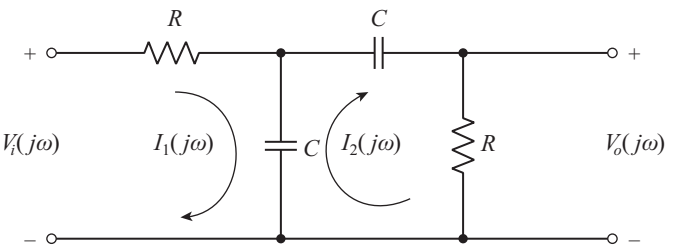

FIGURE 1.2.4. What is $\frac{V_o(j\omega)}{V_i(j\omega)}$?

$$I_1\left(R+\frac{1}{j\omega C}\right)+I_2\left(-\frac{1}{j\omega C}\right)=V_i$$
$$I_1\left(-\frac{1}{j\omega C}\right)+I_2\left(R+\frac{2}{j\omega C}\right)=0.$$

With the 2×2 system determinant D defined as

$$D=\begin{vmatrix}\left(\mathrm{R}+\frac{1}{j\omega C}\right) & \left(-\frac{1}{j\omega C}\right)\\ \left(-\frac{1}{j\omega C}\right) & \left(R+\frac{2}{j\omega C}\right)\end{vmatrix}=R^2-\left(\frac{1}{\omega C}\right)^2-j\frac{3R}{\omega C},$$

Cramer's rule says that

$$I_1=\frac{\begin{vmatrix}V_i & \left(-\frac{1}{j\omega C}\right)\\ 0 & \left(R+\frac{2}{j\omega C}\right)\end{vmatrix}}{D}=V_i(j\omega)\frac{R+\frac{2}{j\omega C}}{R^2-\left(\frac{1}{\omega C}\right)^2-j\frac{3R}{\omega C}},$$

and

$$I_2 = \frac{\begin{vmatrix} \left(\mathrm{R} + \dfrac{1}{\mathrm{j}\omega\mathrm{C}}\right) & V_i \\ \left(-\dfrac{1}{j\omega C}\right) & 0 \end{vmatrix}}{D} = V_i(j\omega)\frac{-j\dfrac{1}{\omega C}}{R^2 - \left(\dfrac{1}{\omega C}\right)^2 - j\dfrac{3R}{\omega C}}.$$

Finally, observing that $V_o(j\omega) = I_2(j\omega)R$, we see that the transfer function is

$$\frac{V_o(j\omega)}{V_i(j\omega)} = H(j\omega) = \frac{-j\dfrac{R}{\omega C}}{R^2 - \left(\dfrac{1}{\omega C}\right)^2 - j\dfrac{3R}{\omega C}}.$$

We can draw one immediate, quite interesting conclusion from this result: $H(j\omega)$ is purely real when the real part of the denominator vanishes. That is, if $\omega = \dfrac{1}{RC}$, then we have $H(j\omega) = +\dfrac{1}{3}$. We won't pursue the implications of this (which are profound) until later in this chapter except to note for now that this property can be used to construct a sinusoidal oscillator. Oscillators are obviously important in radio *transmitters*, but less obvious at this point is that oscillators are also used in radio *receivers*. I'll remind you of Figure 1.2.4 again later in this chapter.

The transfer functions of all but the simplest circuits (for example, all resistors) used in radio will be complex. That is, $H(j\omega)$ will, in general, consist of both amplitude *and* phase response functions, and so

$$H(j\omega) = |H(j\omega)|e^{j\theta(\omega)},$$

where $\theta(\omega)$ is the phase shift that occurs from input to output for a sinusoid at frequency ω. I'll say more about $\theta(\omega)$ later in this chapter.

You'll notice that we did not need $I_1(j\omega)$ to find $H(j\omega)$. Knowledge of $I_1(j\omega)$ nevertheless provides important information. Knowing $I_1(j\omega)$ in terms of $V_i(j\omega)$ allows us to calculate $\dfrac{V_i(j\omega)}{I_1(j\omega)} = Z_i(j\omega)$, the ac

impedance "seen" by the signal source that generates $V_i(j\omega)$. This is important to know because that impedance determines the current the signal source has to be able to provide. So,

$$Z_i(j\omega)=\frac{R^2-\left(\frac{1}{\omega C}\right)^2-j\frac{3R}{\omega C}}{R+\frac{2}{j\omega C}},$$

which, at the frequency $\omega=\frac{1}{RC}$ reduces to

$$Z_i(j\omega)=\frac{-j\frac{3R}{\frac{1}{RC}C}}{R+\frac{2}{j\frac{1}{RC}C}}=\frac{-j3R^2}{R-j2R}=-j\frac{3R}{1-j2}=R\left[-j\frac{3(1+j2)}{(1-j2)(1+j2)}\right]$$
$$=R\left[-j\frac{3+j6}{1+4}\right]=R\left(\frac{6-j3}{5}\right)=R(1.2-j0.6).$$

Thus, while the transfer function of the circuit of Figure 1.2.4 is purely real at $\omega=\frac{1}{RC}$, the input impedance is complex (the negative imaginary part of $Z_i(j\omega)$ means the input impedance "acts like" a capacitor (which, given the components in the circuit, should be no surprise!).

Be particularly careful to notice this important conclusion from our result for $Z_i(j\omega)$: while we can vary either R or C (actually either *both* of the matched R's together or *both* of the matched C's together, because the two resistors are assumed to be equal, and the two capacitors are assumed to be equal[13]) to vary the frequency at which $H(j\omega)$ is purely real, if we choose to vary the two R's we will also vary $Z_i(j\omega)$. If we choose to vary the two C's, however, then we can vary the frequency at which $H(j\omega)$ is purely real while keeping the input impedance fixed. In that case, the V_i signal source "sees" an

13 To simultaneously vary multiple matched-value components, radio engineers use a "ganged shaft" that allows turning a single control-panel knob to rotate a shaft on which all the variable components are mechanically mounted.

unchanging current demand, a property of great importance in building the variable-frequency oscillator circuits we will later encounter in radio receivers.

To end this discussion of transfer functions, let me show you one more thing we can do with them. Suppose we apply a sinusoid at frequency α as the input. That is, suppose $v_i(t) = \sin(\alpha t)$. What is the resulting output $v_o(t)$? From Euler's identity we have

$$v_i(t) = \frac{1}{j2}(e^{j\alpha t} - e^{-j\alpha t}).$$

The output of a circuit is simply the sum of each complex exponential term of $v_i(t)$ multiplied by the transfer function of the circuit *evaluated at the frequency of the input term*.[14] So, for the circuit of Figure 1.2.4,

$$v_o(t) = \frac{1}{j2}\left[e^{j\alpha t}H(j\alpha) - e^{-j\alpha t}H(-j\alpha)\right]$$

$$= \frac{1}{j2}\left[e^{j\alpha t}\frac{-j\frac{R}{\alpha C}}{R^2 - \left(\frac{1}{\alpha C}\right)^2 - j\frac{3R}{\alpha C}} - e^{-j\alpha t}\frac{-j\frac{R}{-\alpha C}}{R^2 - \left(\frac{1}{-\alpha C}\right)^2 - j\frac{3R}{-\alpha C}}\right].$$

There are a lot of j's in this expression, but since the input $v_i(t)$ is real-valued, and since the circuit of Figure 1.2.4 is made from real hardware, we know that $v_o(t)$ *has* to be real, too. Is it? Yes, and you can see that *by inspection* if you write

$$v_o(t) = \frac{1}{j2}\left[e^{j\alpha t}\frac{-j\frac{R}{\alpha C}}{R^2 - \left(\frac{1}{\alpha C}\right)^2 - j\frac{3R}{\alpha C}} - e^{-j\alpha t}\frac{j\frac{R}{\alpha C}}{R^2 - \left(\frac{1}{\alpha C}\right)^2 + j\frac{3R}{\alpha C}}\right].$$

14 Note, carefully, that there are *two* frequencies here: $+\alpha$ and $-\alpha$. The concept of a negative frequency might seem a bit "science fictiony," but we have already encountered a simple physical interpretation, namely, the *clockwise*-rotating vector in the box that opens this section. Later in the book, when we get to the sidebands of a modulated rf carrier wave, in both AM and FM radio, you'll see that "negative" frequencies possess an undeniable physical reality.

The expression inside the square brackets is the difference of conjugates and so is equal to the imaginary part of the first term times $j2$ (which is cancelled by the $\frac{1}{j2}$ in front of the brackets). Specifically,

$$v_o(t) = \text{Im}\left[e^{j\alpha t} \frac{-j\frac{R}{\alpha C}}{R^2 - \left(\frac{1}{\alpha C}\right)^2 - j\frac{3R}{\alpha C}} \right]$$

$$= \text{Im}\left[e^{j\alpha t} \frac{-j\frac{R}{\alpha C}\left\{R^2 - \left(\frac{1}{\alpha C}\right)^2 + j\frac{3R}{\alpha C}\right\}}{\left\{R^2 - \left(\frac{1}{\alpha C}\right)^2\right\}^2 + \left(\frac{3R}{\alpha C}\right)^2} \right],$$

or

$$v_o(t) = \text{Im}\left[\frac{\{\cos(\alpha t) + j\ \sin(\alpha t)\}\left\{-j\frac{R}{\alpha C}\left[R^2 - \left(\frac{1}{\alpha C}\right)^2 + j\frac{3R}{\alpha C}\right]\right\}}{\left\{R^2 - \left(\frac{1}{\alpha C}\right)^2\right\}^2 + \left(\frac{3R}{\alpha C}\right)^2} \right].$$

Thus,

$$v_o(t) = \frac{\frac{3R^2}{(\alpha C)^2}\sin(\alpha t) - \frac{R}{\alpha C}\left\{R^2 - \left(\frac{1}{\alpha C}\right)^2\right\}\cos(\alpha t)}{\left\{R^2 - \left(\frac{1}{\alpha C}\right)^2\right\}^2 + \left(\frac{3R}{\alpha C}\right)^2}.$$

As a partial check on our calculations, notice that if $\alpha = \frac{1}{RC}$, then this expression reduces to $v_o(t) = \frac{1}{3}\sin\left(\frac{t}{RC}\right)$ when $v_i(t) = \sin\left(\frac{t}{RC}\right)$, giving $H\left(j\frac{1}{RC}\right) = +\frac{1}{3}$, which we found earlier.

1.3 Van der Pol's Negative Resistance Oscillator Equation

Almost from the invention of the electric battery in 1799–1800, by the Italian physical chemist Alessandro Volta (1745–1827), it was known that a low-voltage source (a few tens of volts) able to supply a continuous large current (hundreds of amperes) could generate an electrical arc of intense brilliance.[15] That is, if two electrodes in contact, carrying this current, are slowly pulled apart to form a gap, the electrode current can continue to flow across the gap, appearing as a *flame* of ionized atmospheric gases and vaporized electrode material. By the 1890s it was known that a plot of the gap current versus the gap voltage drop had the surprising behavior shown in Figure 1.3.1.

The surprising feature of Figure 1.3.1 is, of course, that the current-voltage curve of the electric arc has a *kink*, that is, an interval where a *decrease* in gap current is associated with an *increase* in gap voltage drop, behavior certainly not at all like the Ohm's law linear behavior of a resistor. The total voltage drop divided by the total gap current is always positive, but *in the kink* the *dynamic* ratio $\frac{di}{dv}$ is negative (the slope of the kink is negative), and for that reason the electric arc was said to have a negative ac resistance. With some very clever engineering, this feature of the electric arc was used to neutralize the energy-dissipating positive R of FitzGerald's oscillating circuit, allowing the construction of very powerful radio transmitters. Arc radio[16] itself is of only historical interest today, but the mathematical theory of negative resistance is of continuing interest, as it also appears in the electronics of modern radio.

15 The Cornish chemist Humphry Davy (1778–1829), mentor to the young Faraday, invented the arc lamp in 1809.

16 Not to be confused with the biblical Ark of the Covenant, said (in the first Indiana Jones movie, the 1981 *Raiders of the Lost Ark*) to be "a radio for speaking to God." The Ark, built by Moses according to detailed instructions from God (Exodus 25) to hold the stone tablets of the Ten Commandments, is described in various ancient Jewish legends as being surrounded by sparks and so was perhaps electrical in nature. Further, when Uzzah touched the Ark (2 Samuel 6:67) he instantly died (electrocuted?). In Exodus 25:22 the Lord tells Moses he will speak to him from the Ark, and this was the motivation for the movie line claiming the Ark to be a radio. Well, I have to admit it's a thought-provoking assertion, but we'll find *our* inspiration for radio in this book to be more from Maxwell's *Treatise* than we will from the Bible.

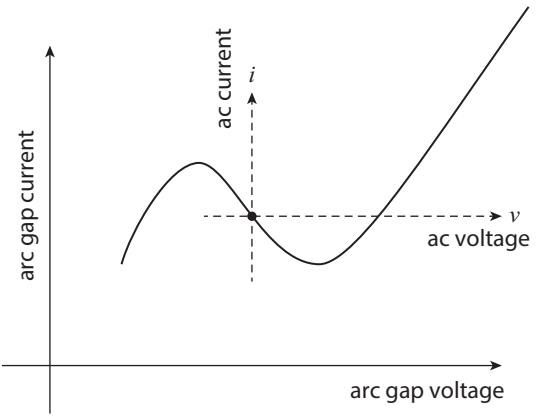

FIGURE 1.3.1. Current versus voltage for an electric arc.

In Figure 1.3.2 we see FitzGerald's circuit connected in parallel with an arc which is powered by a low-voltage, high-current dc energy source (which is itself in series with an inductor called the *choke coil*, a name that will be explained in just a moment). To understand what is happening in this enhanced FitzGerald circuit, you have to visualize two distinct current loops. First, there is the dc loop formed by the energy source, the choke coil (with a small ohmic resistance important at dc but presenting a relatively high ac impedance at the frequency at which the circuit oscillates), and the arc. Second, there is an ac current loop formed by the R, L, C, and again, the arc. The total ac resistance in this second loop is the sum of R and the dynamic ac resistance of the arc (which, being negative, can result in a net ac resistance of zero). Because of the choke coil, the oscillations in the ac loop cannot "leak back" through the dc source, which typically has a very low resistance. Such leakage would result in energy loss via heating of the dc source.

The arc current consists, then, of two components: a large, steady dc current, on top of which is superimposed an oscillating (that is, an ac) component. This is indicated in Figure 1.3.1 by the dashed axes centered on the midpoint of the negative resistance kink. If we imagine the arc operates at that midpoint ($i = v = 0$) when there are no oscillations, then when we say the circuit of Figure 1.3.2 oscillates, we mean we are interested in the current/voltage *deviations* around

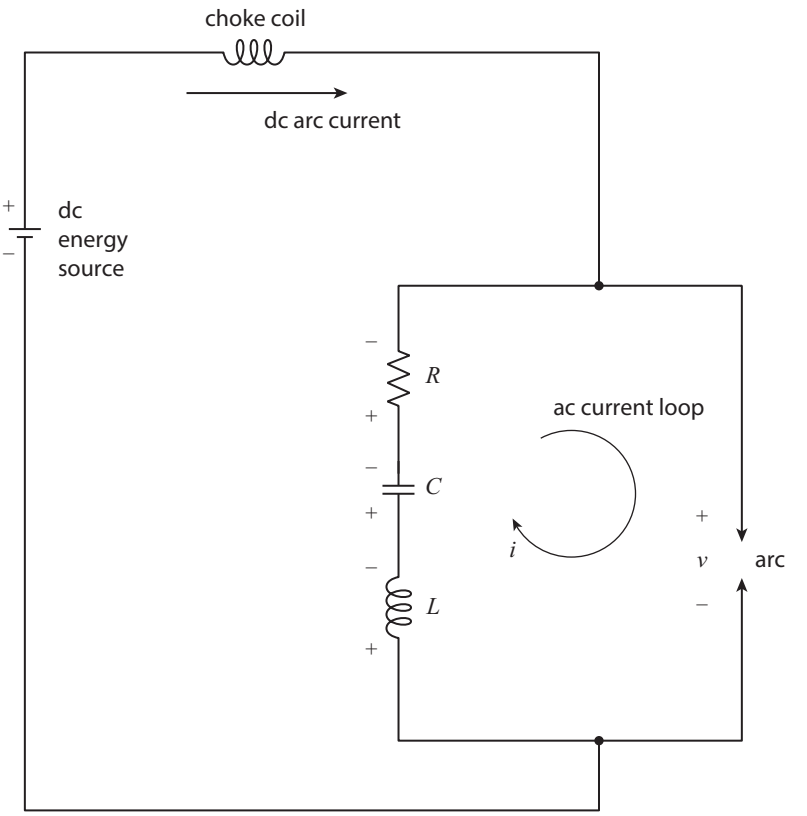

FIGURE 1.3.2. The ac current and voltage (i, v) in FitzGerald's series circuit, in parallel with an electric arc.

that dc midpoint. Now, be sure to understand that all this business about the electric arc *itself* is not the central issue. What is important is the *negative resistance* of the arc; I use the arc here simply to give you a physical model.to envision. When we get to the superregenerative radio receiver in chapter 3, you'll again see negative resistance mentioned in connection with oscillatory behavior.

In any case, such arc-enhanced versions of FitzGerald's circuit were unable to oscillate at frequencies beyond about 60 kHz or so,[17]

17 Because of engineering difficulties that are discussed in *The Science of Radio* (note 8), none of which I'll pursue here because Hardy couldn't have cared less about such things. If I really wanted to drive Hardy into a coma, I could next tell him that these same difficulties (and their solutions) appear in the physics of circuit breakers with superfast tripping times. But I wouldn't actually want to do that, and so I won't do it here, either.

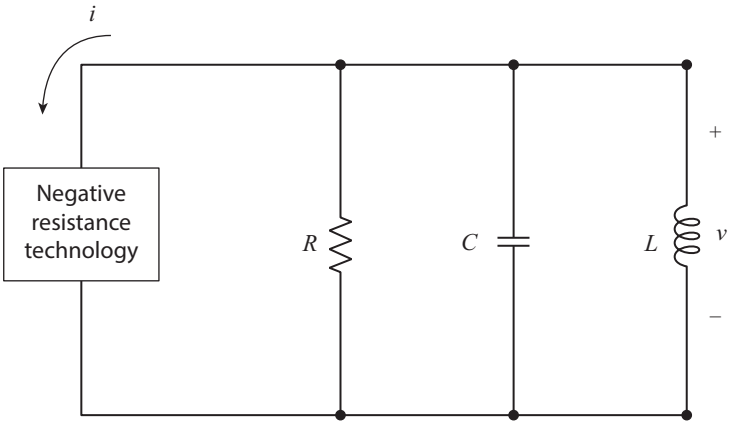

FIGURE 1.3.3. Van der Pol's parallel version of FitzGerald's oscillating series circuit.

and by the early 1920s the use of arc transmitters in radio was commercially dead. Van der Pol, however (see the box in the preface) studied a mathematically equivalent[18] parallel version (as shown in Figure 1.3.3) of FitzGerald's series circuit while retaining the idea of a negative resistance kink in the ac voltage/current behavior of whatever nonlinear technology (present or future) was under study.

As stated in note 18, from Kirchhoff's current law Van der Pol immediately wrote

$$i+\frac{v}{R}+C\frac{dv}{dt}+\frac{1}{L}\int v\ dt=0,$$

and he modeled the kink in the negative resistance technology box at the far left of Figure 1.3.3 with the equation

$$i=-av+bv^3,$$

18 Here's what *mathematically equivalent* means. If we write Kirchhoff's voltage law for the ac loop in Figure 1.3.2, we have $v+L\frac{di}{dt}+\frac{1}{C}\int i\ dt+iR=0$.

If we write Kirchhoff's current law for Figure 1.3.3, we have $i+\frac{v}{R}+C\frac{dv}{dt}+\frac{1}{L}\int v\ dt=0$, which is the voltage loop equation with v and i swapped, with R and $1/R$ swapped, and with L and C swapped. With those trivial symbol changes, FitzGerald's original *series* ac circuit becomes Van der Pol's *parallel* ac circuit. Electrical engineers say that each circuit is the *dual* of the other; the behavior of i in Figure 1.3.2 is the behavior of v in Figure 1.3.3.

where a and b are positive constants. Differentiating both equations, we get

$$\frac{di}{dt}+\frac{1}{R}\frac{dv}{dt}+C\frac{d^2v}{dt^2}+\frac{1}{L}v=0,$$

and

$$\frac{di}{dt}=-a\frac{dv}{dt}+3bv^2\frac{dv}{dt}.$$

I'll leave it to you to confirm that substituting the second equation into the first one and doing a little algebra gives the result

$$LC\frac{d^2v}{dt^2}+\left[L\left(\frac{1}{R}-a\right)+3bLv^2\right]\frac{dv}{dt}+v=0.$$

Next, we change the variable to $x=\omega_0 t$, where $\omega_0=\dfrac{1}{\sqrt{LC}}$. Then, $dx=\omega_0 dt$, or $dt=\dfrac{dx}{\omega_0}$, and so

$$\frac{dv}{dt}=\frac{dv}{\dfrac{dx}{\omega_0}}=\omega_0\frac{dv}{dx}.$$

Thus,

$$\frac{d^2v}{dt^2}=\frac{d}{dt}\left(\frac{dv}{dt}\right)=\frac{d}{dt}\left(\omega_0\frac{dv}{dx}\right)=\frac{d}{\dfrac{dx}{\omega_0}}\left(\omega_0\frac{dv}{dx}\right)=\omega_0\frac{d}{dx}\left(\omega_0\frac{dv}{dx}\right)$$

$$=\omega_0{}^2\frac{d^2v}{dx^2}=\frac{1}{LC}\frac{d^2v}{dx^2}.$$

Substituting these two results for $\dfrac{dv}{dt}$ and $\dfrac{d^2v}{dt^2}$ into the equation just before we change the variable to x, we have

$$\frac{d^2v}{dx^2}+L\left[\left(\frac{1}{R}-a\right)+3bv^2\right]\omega_0\frac{dv}{dx}+v=0,$$

or after just two or three more easy steps, we arrive at

$$\frac{d^2v}{dx^2}-\left[\sqrt{\frac{L}{C}}\left(\frac{1}{R}-a\right)-3b\sqrt{\frac{L}{C}}+v^2\right]\frac{dv}{dx}+v=0.$$

Now, to finish, we write $\varepsilon=\sqrt{\frac{L}{C}}\left(\frac{1}{R}-a\right)$ and make the change of variable $v=hu$, where h is the constant such that $h^2=\frac{\varepsilon}{3b\sqrt{\frac{L}{C}}}$. In just a couple more easy steps of algebra we arrive at Van der Pol's equation:

$$\frac{d^2u}{dx^2}-\varepsilon(1-u^2)\frac{du}{dx}+u=0.$$

In this equation u is a normalized v as a function of x (which, in turn, is a normalized time). The parameter ε has absorbed the values of a, R, L, and C, while h (which is an amplitude-scaling parameter relating u and v) has absorbed b. Van der Pol's nonlinear differential equation is not "easy" to solve, and he was able to find analytical solutions only for the case of $\varepsilon \ll 1$, for which he found the remarkable result that the solutions are periodic with a normalized amplitude of 2. In *The Science of Radio* (note 8) I work through, in detail, how Van der Pol did this. It's elementary, but pretty tricky. Van der Pol was a very clever engineering analyst (and I think even Hardy would have concluded that).

Hardy's friend Littlewood came to his study of Van der Pol's equation in response to a January 1938 memorandum from the British Radio Research Board asking for "really expert guidance" from pure mathematicians in helping engineers understand the behavior of "certain types of non-linear differential equations involved in radio engineering." A copy of the memorandum was sent to the London Mathematical Society, where it caught the eye of Mary Cartwright (1900–1998), an English mathematician who started her doctoral studies at Oxford under Hardy (but finished with a different thesis advisor when Hardy left Oxford for a sabbatical leave at Princeton University and Caltech, 1928–1929). Cartwright knew Littlewood (who had, in June 1930, traveled to Oxford to supervise her doctoral examination),

she got him interested in Van der Pol's equation, and the two of them decided to jointly respond to the 1938 memorandum.[19]

Their starting point was a 1934 paper published by Van der Pol that included "graphically integrated" solutions for various values of ε. Littlewood was able to show that the oscillation amplitude was *not* exactly 2 for ε "small," but later Cartwright and Littlewood further showed that as $\varepsilon \to \infty$ the oscillation amplitude *did* approach 2 from above. With a modern home computer and powerful software (I use MATLAB), it is today easy to confirm these results. Figure 1.3.4, for example, shows a computer solution for one of the values of ε in Van der Pol's paper,[20] and it is virtually identical with Van der Pol's graphical solution (see the following box for how this figure was created). To quote from Van der Pol's paper, the solution "represents the slow building up of an approximately sinusoidal oscillation." The final amplitude of those oscillations does appear to be pretty close to 2. For $\varepsilon \gg 1$ the oscillations are decidedly *not* sinusoidal.

The computer-generated solution to Van der Pol's differential equation was obtained by the standard method of writing an nth-order differential equation as a system of n first-order differential equations. That is, we start by defining

$$u_1(x) = u(x), \quad u_2(x) = \frac{du_1}{dx}.$$

Then,

$$\frac{du_1}{dx} = \frac{du}{dx} = u_2, \quad \frac{du_2}{dx} = \frac{d^2u_1}{dx^2} = \frac{d^2u}{dx^2} = (1 - u_1^2)u_2 - u_1.$$

19 See Shawnee L. McMurran and James J. Tattersall, "The Mathematical Collaboration of M. L. Cartwright and J. E. Littlewood," *American Mathematical Monthly*, December 1996, pp. 837–845.

20 B. van der Pol, "The Nonlinear Theory of Electric Oscillations," *Proceedings of the IRE* (Institute of Radio Engineers), September 1934, pp. 1051–1086.

(continued...)

Index