

This text has been transcribed, without permission, from:

John Strawn, ed. 1985. *Digital audio signal processing: An anthology* Los Altos, Calif.:Kaufmann.

[More information about this text](#)

SPIRAL SYNTHESIS

Tracy Lind Petersen

WHY SPIRALS?

Spiral formations are known to occur in the structure of proteins and in the great arms of the Andromeda nebula. Spirals twist their way through an amazing variety of things, including seashells, plants, and DNA. Because of the structure of their eyes, even certain insects follow the path of a logarithmic spiral as they make their way into the candle flame.

It is also true that acoustic energy is ultimately converted to neural impulses within the auditory nerve at the location of the inner ear, which has a spiral geometry rather like that of the seashell. Perhaps by coincidence, when the frequency sensitivity of the ear is given a signal processing interpretation, the z transform representation of this sensitivity places simple resonances along a spiral curve in the z transform domain. The time-waveforms which emerge as responses (*impulse responses*, in the language of signal processing) from a given resonance also turn out to be complex spirals which we will soon discuss in greater detail.

We might speculate, since nature has had some time to invoke natural selection in the synthesis of these microtransducers we call our ears, that we have in some sense a very special mechanism with which to decode the acoustic world. Indeed, the spiral chamber of the inner ear (cochlea) is of essentially the same structure, within a scale factor, in cats, humans, and elephants ([Greenwood 1962](#)). It is sometimes stated that the ear performs Fourier frequency analysis. In fact, the peripheral auditory system performs a type of filter-bank analysis with a frequency-dependent bandwidth characteristic ([Scharf 1961](#)) which differs considerably from short-time Fourier analysis. These measured auditory bandwidths are of such a fundamental nature that they are called *critical bands*. The *critical band transform* (CBT) ([Petersen 1980](#); [Petersen and Boll 1983](#)) facilitates the representation of acoustic signals in terms of auditory critical bandwidth parameters. CBT analysis allows us to view acoustic signals in terms of complex spirals which are easily synthesized with a simple recursive technique. Furthermore, these spirals are naturally represented in terms of quadrature sinusoids modulated in both amplitude and phase. While the development of this spiral synthesis technique proceeds from the theory of z transforms, the method will be presented with an emphasis on that which is intuitive.

MOVING THE SPIRAL

If we stretch a helical spring parallel to the wall and floor of a room and project its (stationary) shadow on the wall, we see a sine wave. If, without moving the spring, we project its shadow on the floor, we see a second sine wave which is out of phase with the one on the wall by $\pm 90^\circ$. While this is an idealized picture, if we accept the spring shadow on the wall as the function $\cos(t)$ and its projection on the floor as $\sin(t)$, then the sinusoid on the wall is said to be *in quadrature* with the one on the floor. [Figure 3.1](#) is an illustration of two such waveforms, projected from a single spiral. In the language of signal processing, we are talking about an analytic signal where the real part is projected on the "wall" (labeled Re) and the imaginary part is projected on the "floor" (labeled Im).

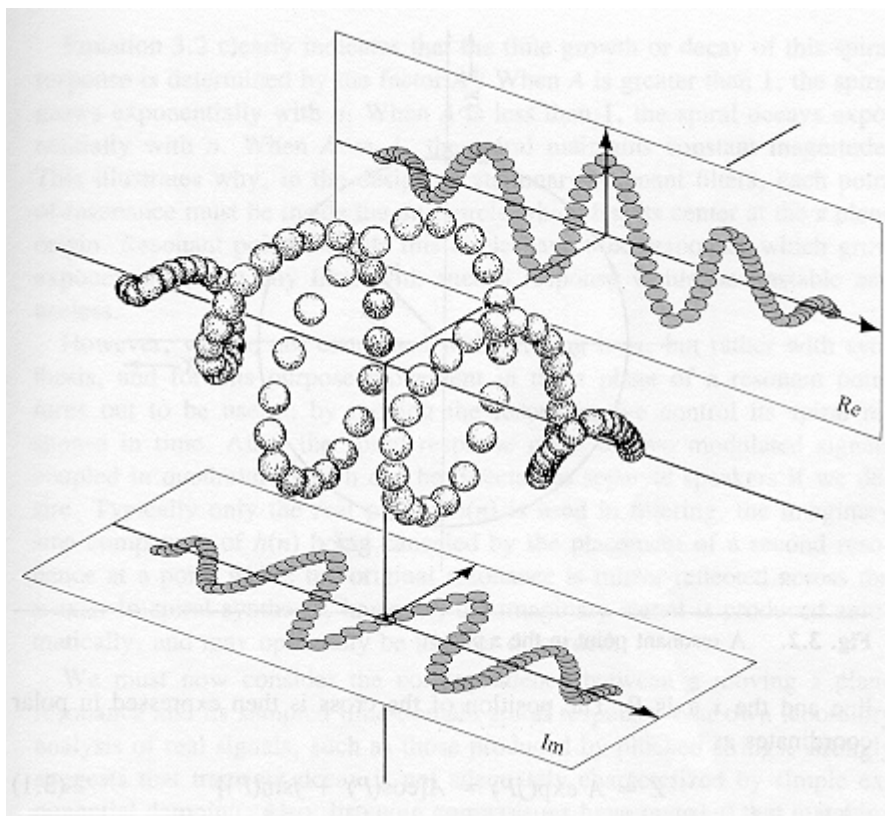


Fig. 3.1. Real and imaginary projections of a complex spiral

The purpose of this visual analogy is to show that sinusoids are representable as projections of a spiral in three-dimensional space onto a flat surface. Mathematically, sine waves may be expressed in terms of this spiral. Our synthesis technique will allow us to control the evolution of such a spiral as it is being generated.

The graphical representation of an acoustic waveform gives us a picture of sound in time. Our shadow projections in the previous example are such pictures. It is often desirable to represent a given sound in terms of its frequency content as well, and such frequency "pictures" may be obtained through an application of, for example, Fourier, Laplace, or critical band transform techniques. When we work with sampled waveforms, the z transform domain is very useful. Sampled time-waveforms have corresponding z transform representations which reveal their frequency content. Also, certain classes of filters have convenient representations in the z transform domain. We refer to this domain as the z plane.

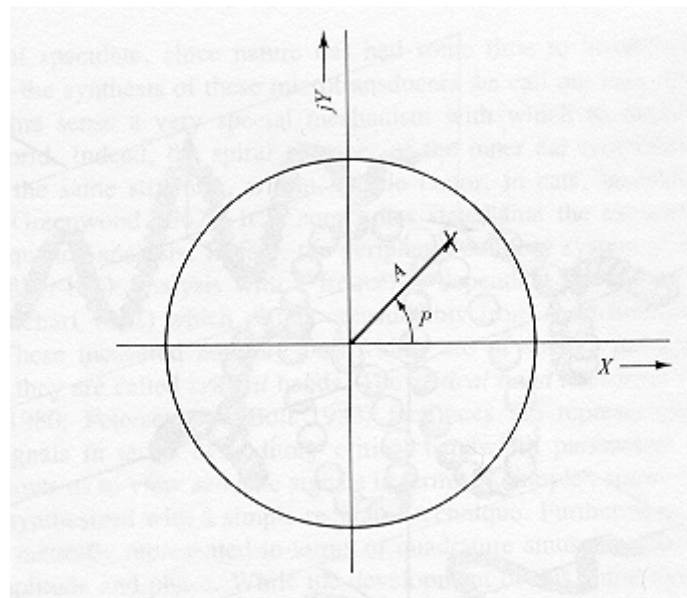


Fig. 3.2. A resonant point in the z plane

The z plane allows us to construct a simple picture which will help to explain our synthesis technique intuitively. The picture also allows us to introduce a basic mathematical concept which underlies the approach. We will begin by describing the z plane picture of a simple resonant filter. [Figure 3.2](#) shows a resonant point, or pole, marked by a cross in the z plane. For our purposes, the important part of this picture is given by the position of the cross, as its z plane

location directly affects what kind of sampled waveform is being generated in the time space. We define the cross (resonant) position by drawing a line which connects it with the z plane origin. We call the length of this line A , and we call the radian angle between this line and the x axis P . The position of the cross is then expressed in polar coordinates as

$$Z = A \exp(jP) = A[\cos(P) + j\sin(P)] \quad (3.1)$$

Given that a filter has this particular resonant point, the time response of this filter when it receives a single input pulse will, according to z transform theory, be a spiral which proceeds down the time line. Its sinusoidal projection (on the floor in our previous analogy) will be given by the samples $A^n \sin(nP)$ and its cosine projection (on the wall) will be given by the samples $A^n \cos(nP)$, where time now is represented by uniformly spaced integer values of the discrete variable n . Because the sine projection (on the floor) is the imaginary part of our complex time response, we can show by Euler's identity that our samples of this spiral may be represented as

$$\begin{aligned} h(n) &= A^n [\cos(nP) + j\sin(nP)] \\ &= A^n [\exp(jnP)] \\ &= [A \exp(jP)]^n \end{aligned} \quad (3.2)$$

Referring back to [equation 3.1](#), we may now conclude that if the z plane resonance occurs at $z = A \exp(jP)$, the time response to a unit sample input is $[A \exp(jP)]^n$, and, in general, a resonance at $z = q$ produces a unit sample response q^n , where n assumes successive integer values in time.

[Equation 3.2](#) clearly indicates that the time growth or decay of this spiral response is determined by the factor A^n : When A is greater than 1, the spiral grows exponentially with n . When A is less than 1, the spiral decays exponentially with n . When $A = 1$, the spiral maintains constant magnitude. This illustrates why, in the design of stationary resonant filters, each point of resonance must be inside the unit circle which has its center at the z plane origin. Resonant points outside this circle have time responses which grow exponentially, and any filter with such a response would be unstable and useless.

However, we are not concerned with filtering here, but rather with synthesis, and for this purpose movement in the z plane of a resonant point turns out to be useful: by moving the resonance we control its spiral response in time. Also, the spiral response gives us two modulated signals coupled in quadrature which can be directed to separate speakers if we desire. Typically only the real part of $h(n)$ is used in filtering, the imaginary sine component of $h(n)$ being canceled by the placement of a second resonance at a point where the original resonance is mirror-reflected across the x axis. In spiral synthesis, however, the imaginary signal is produced automatically, and may optionally be used or discarded.

We must now consider the correspondence between a moving z plane resonance and its sampled time-domain spiral response. Our own laboratory analysis of real signals, such as those produced by plucked strings, strongly suggests that transient decay is not adequately characterized by simple exponential damping. Also, listening comparisons have revealed that matching real transient responses with as many as 90 poles at a 16 kHz sample rate does not produce a convincing imitation of the original sound quality. Particularly, an instrument such as the plucked monochord has a long transient which is not easily modeled as simple exponential decay. Spiral synthesis allows the manipulation of transient decay characteristics. We have only to excite the system once, with a single pulse, and the response can be made to ring on in a continuously evolving fashion, by allowing the point of resonance to drift in and out of the z plane unit circle.

The angle P in [figure 3.2](#) appears as an argument to the sine and cosine terms in [equation 3.2](#). From this we see that P is the number of radians per sample in the complex time response $h(n)$. For a fixed sample frequency, then, P sets the frequency of our synthesized time response where the conditions of the sampling theorem require that P be less than π radians per sample for any given frequency.

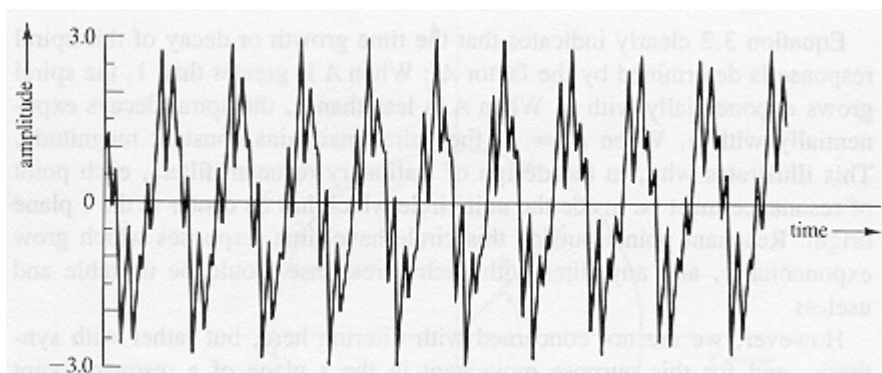


Fig. 3.3. Real part of a signal generated using the program in [code listing 3.1](#) where the number of samples was $n_{\text{cnt}} = 100$, radians per sample of the carrier was $\text{angl} = \pi/100$, initial pole distance was $\text{mag} = 1.01$, modulating amplitude was $\text{modsc1} = 0.2$, and radians per sample of the modulator was $\text{phzdel} = \pi/17$.

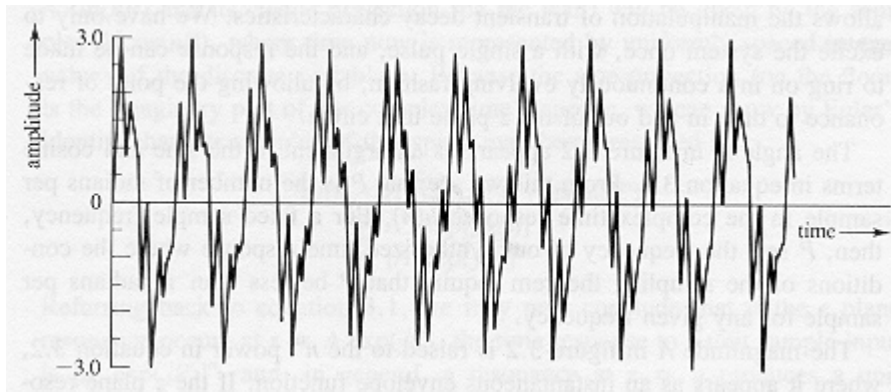


Fig. 3.4. Imaginary part of the signal shown in [figure 3.3](#)

The magnitude A in [figure 3.2](#) is raised to the n^{th} power in [equation 3.2](#), where it appears as an instantaneous envelope function. If the z plane resonance becomes dynamic, both A and P become functions of time variable n , and we accommodate this in our notation by writing the magnitude factor as $A(n)$ and the phase term as $P(n)$.

Z transform theory allows us to write a recursive expression which generates successive samples of the spiral response as we proceed in time. If we let $q(n) = A(n)\exp[P(n)]$ be the z plane resonance, then it can be shown that an expression which generates the modulated spiral response is

$$h(n) = \delta(n) + q(n)h(n-1), \quad h(n) = 0, \quad n < 0, \quad (3.3)$$

where $\delta(n)$, the unit sample function, is 1 when $n = 0$ and 0 otherwise.

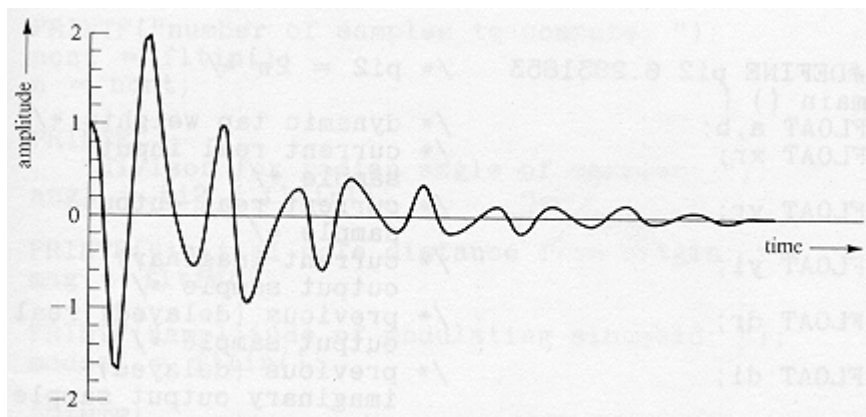


Fig. 3.5. Real part of a second spiral generated using the program in [code listing 3.1](#) but with different parameter values. Here $n_{\text{cnt}} = 160$, $\text{angl} = \pi/16$, $\text{mag} = 0.98$, $\text{modsc1} = 0.15$, and $\text{phzdel} = \pi/23$.

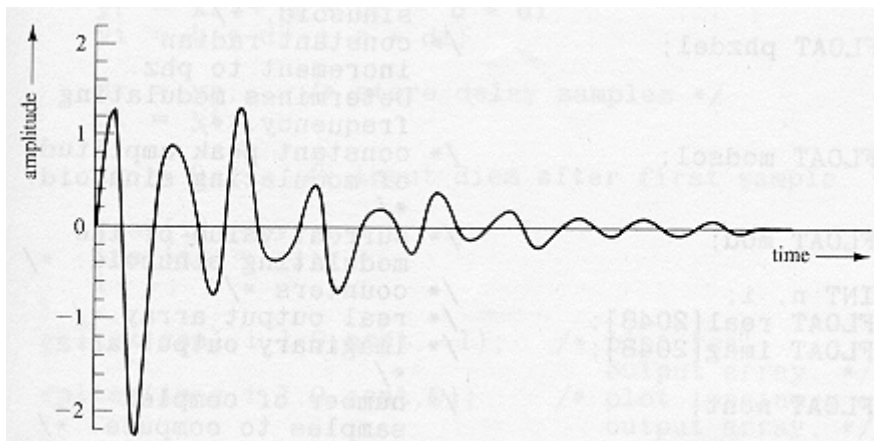


Fig. 3.6. Imaginary part of the signal shown in [figure 3.5](#)

The capability to modulate real and imaginary output signals should now be clear. When $A(n)$ is greater than 1, the amplitude of the time response starts to increase. When $A(n)$ is less than 1, the time response will begin to decay. When $P(n)$ is fixed, the frequency of the synthesized sine and cosine will be a constant $P(n)$ radians per sample. When $P(n)$ changes in time, the instantaneous phase of the synthesized sine and cosine are modulated accordingly. The current value of the phase will be the accumulated sum of all $P(n)$ and the current value of the signal envelope will be the accumulated product of all $A(n)$. Thus, at time n the process is bounded by

$$\prod_{i=1}^n A(i).$$

CODE LISTING 3.1 C Implementation of Spiral Synthesis

```

#define pi2 6.2831853 /* pi2 = 2π */
main() {
  FLOAT a,b;          /* dynamic tap weights */
  FLOAT xr;           /* current real input
                      sample */
  FLOAT yr;           /* current real output
                      sample */
  FLOAT yi;           /* current imaginary
                      output sample */
  FLOAT dr;           /* previous (delayed) real
                      output sample */
  FLOAT di;           /* previous (delayed)
                      imaginary output sample
                      */
  FLOAT angl;        /* radians per sample of
                      output sinusoid before
                      modulation. (Sets
                      carrier frequency.) */
  FLOAT angl;        /* current radians per
                      sample of output
                      sinusoid after
                      modulation. */
  FLOAT mag;         /* initial distance of
                      pole position from z
                      plane origin. */
  FLOAT phz;         /* current radian angle of
                      the modulating
                      sinusoid. */
  FLOAT phzdel;      /* constant radian
                      increment to phz.
                      Determines modulating
                      frequency. */
  FLOAT modscl;      /* constant peak amplitude
                      of modulating sinusoid.
                      */
  FLOAT mod;         /* current value of the
                      modulating sinusoid. */
  INT n, i;          /* counters */
  FLOAT real[2048];  /* real output array */
  FLOAT imag[2048]; /* imaginary output array
                      */
  FLOAT ncnt;        /* number of complex
                      samples to compute. */
  DOUBLE fltin();    /* inputs floating-point
                      number from tty. */
  DOUBLE SIN(),COS();

  PRINTF("number of samples to compute: ");
  ncnt = fltin();
  n = ncnt;

  PRINTF(
    "divisor for radian angle of carrier: ");
  angl = pi2 / fltin();

```

```

PRINTF("initial pole distance from origin: ");
mag = fltin();

PRINTF(
    "two-pi divisor for radian angle of modulator: ");
phzdel = pi2 / fltin();

phz = 0.;
dr = di = 0.;    /* initialize delays to zero. */
xr = 1.0;       /* initialize input with a
                unit pulse. */

i = 0;
WHILE(n--) {    /* do it! */
mod = modscl * SIN(phz);
phz = phz + phzdel;
mg2 = mag + mod;
angl2 = angl1 + mod;

a = mg2 * COS(angl2); /* set current value*/
b = mg2 * SIN(angl2); /* of tap weights. */

    yr = xr + a * dr - b * di;
    yi = b * dr + a * di;

    dr = yr; /* store delay samples */
    di = yi;

    xr = 0.; /* input dies after first sample. */

    real[i] = yr;
    imag[i] = yi;
    i++;
}
rpltek(real,i,1.0,ncnt,-1); /* plot real
                             output array. */
rpltek(imag,i,1.0,ncnt,0); /* plot imaginary
                             output array. */
}

```

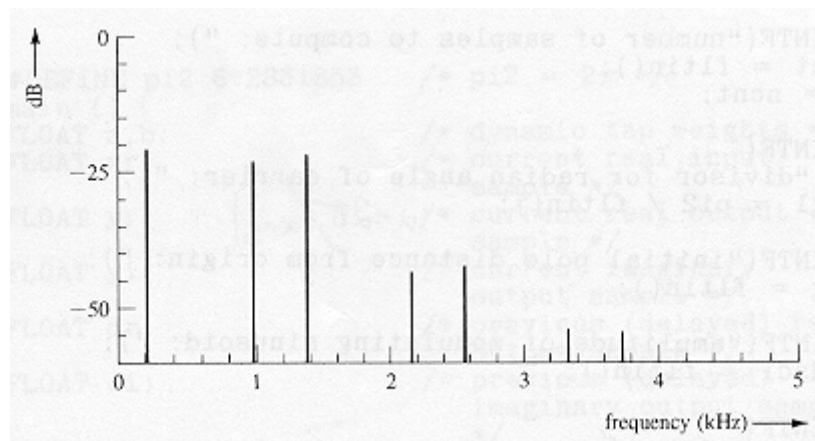


Fig. 3.7. Spectrum of the signal shown in [figure 3.3](#), generated at a sampling rate of 20 kHz. The carrier occurs at $20\,000/100 = 200$ Hz. Sidebands are generated at $f_c \pm nf_m$, $n = 1, 2, 3, \dots$, with the modulator $f_m = 20\,000/17 = 1176.5$ Hz. (Figures [3.7](#) and [3.8](#) were supplied by the editor; responsibility for errors in these two figures and the accompanying captions lies with him and not with the author.)

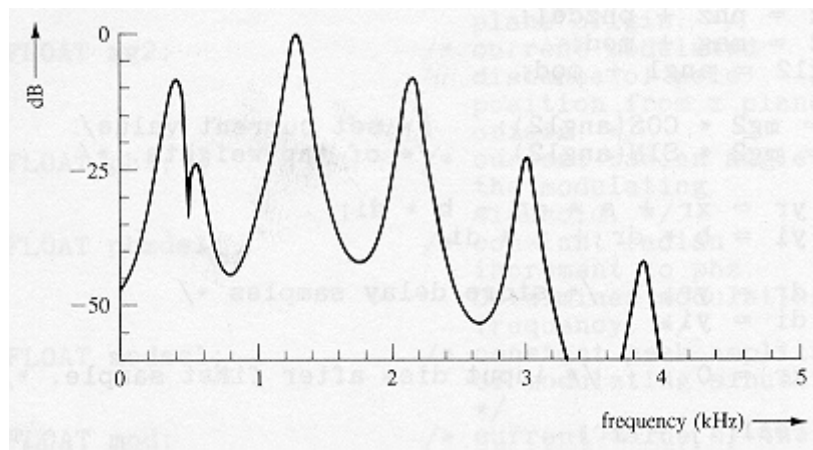


Fig. 3.8. Spectrum of the signal shown in [figure 3.5](#), generated at a sampling rate of 20 kHz. Since the signal is much shorter than that of [figure 3.3](#), the spectral peaks are much broader here. Again, the carrier lies at $20\,000/16 = 1250$ Hz; the modulator is $20\,000/23 = 870$ Hz. The double peak at the left occurs because two sidebands lie very close to each other. The first-order sideband ($1250 - 870$) appears at 380 Hz, and the next-lower sideband ($1250 - 2 \times 870 = -490$) undergoes "0 Hz wraparound" to fall at 490 Hz.

A SYNTHESIS PROGRAM

For purposes of illustration, a generating program is given in [code listing 3.1](#), along with plots of synthesized waveforms. We must remember that the multiplication in [equation 3.3](#) is complex. Two storage locations are required for the delayed sample of $h(n)$ (`dr`, `di` in [code listing 3.1](#)). Current real and imaginary values of $q(n)$ are the dynamic tap weights `a`, `b` in that code listing. The synthesized output which we have so far called $h(n)$ is stored in the arrays `real` and `imag`. A sinusoidal modulator is used; for the sake of simplicity, the same modulation is applied to both the magnitude and the phase of the z plane resonance.

Figures [3.3](#) and [3.4](#) show the real and imaginary parts, respectively, of a complex spiral generated by the program in [code listing 3.1](#). Figures [3.5](#) and [3.6](#) again show real and imaginary parts of a second spiral generated from different input parameter values.

The waveforms we have generated for illustration help to convey the fact that even the simplest approach to modulation can produce time-waveforms which look interesting. However, the independence of amplitude and phase modulators is easily accommodated and allows fine control over the timbral evolution of the spiral.

REFERENCES

- Greenwood, D. D. 1962. Approximate calculation of the dimensions of traveling-wave envelopes in four species. *Journal of the Acoustical Society of America* 34:1364-69.
- Petersen, Tracy L. 1980. Acoustic signal processing in the context of a perceptual model. Technical Report UTEC-CSc-80-113. Ph.D. diss., Department of Computer Science, University of Utah.
- Petersen, Tracy L., and S. F. Boll. 1983. Critical band analysis-synthesis. *IEEE proceedings on Acoustics, Speech, and Signal Processing* ASSP-31(3):656-63.
- Scharf, E. 1961. Complex sounds and critical bands. *Psychological Bulletin* 58:205-17.