Notes on Integration for Digital Data

Herman Jaramillo

May 10, 2016

Contents

Chapter 1

Numerical Integration Algorithms

1.1 Basic Facts

This document assumes that the reader if familiar with Fourier analysis and distribution theory. The integration here is assumed to be performed in uniformly sampled data and, when not explicitly writen, the sampling rate is assume to be the unit $\Delta t = 1$.

Let us start with a note on Fourier transforms. We assume that the forward/inverse temporal Fourier transforms applied to a real function $f(t)$ are given by

$$\begin{aligned} F(\omega) &= \int_{-\infty}^{\infty} f(t) \mathrm{e}^{\mathrm{i}\omega t} dt \\ f(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \mathrm{e}^{-\mathrm{i}\omega t} d\omega. \end{aligned} \tag{1.1.1}$$

With this convention, it should be clear that differentiation in equation 1.1.1 brings a $-\mathrm{i}\omega$ in front of the operator, while integration brings a $-1/(\mathrm{i}\omega) = \mathrm{i}/\omega$. This last factor will be seen often on this document. Different from physicists, electrical engineers consider the opposite sign on the phase $\mathrm{i}\omega t$.

Jon Claerbout has an interesting note in his Seismic Exploration Project (SEP) website [1] about the sign convention in Fourier transform, where he states the convenience of convention under different circumstances.

In addition, for the Z transform, electrical engineers use z^{-1} while physicist use z, for $z = \exp \mathrm{i}\omega\Delta t$. We will go back to this observation below.

The causal integral operator, acting on a function of real variable f, is defined as

$$F(t) = \int_{-\infty}^{t} f(\tau) d\tau. \tag{1.1.2}$$

In filter theory we want to understand convolutional filters based on their impulse responses. A perfect impulse at some time t_0 in the continuum is a Dirac delta distribution

[1] http://sepwww.stanford.edu/theses/sep16/16_29.pdf

$\delta(t - t_0)$. The response to this delta under integration is

$$H(t - t_0) = \int_{-\infty}^{t} \delta(\tau - t_0) d\tau = \begin{cases} 1 & \text{for} \quad t > t_0 \\ 0 & \text{for} \quad t < t_0 \end{cases}$$

Sometimes $H(0) = 0$ or $H(0) = 1$ or $H(0) = 1/2$. For example if we require right continuity we assume $H(0) = 1$, if we require left continuity $H(0) = 0$, but we can also define

$$H(t) = \frac{1}{2}(1 + \operatorname{sgn} t),$$

for which $H(0) = 1/2$. This last definition is useful to find the Fourier transform of $H(t)$, and as we will show, it is the impulse response of the trapezoidal rule integrator.

The Fourier transform of $H(t)$ is given by

$$\mathcal{F}[H(t)] = \frac{1}{2}\,\delta(\omega) + \frac{i}{\omega}. \tag{1.1.3}$$

This is the spectrum of the ideal integrator. The phase shift of $\pi/2$, from

$$\mathrm{i} = \mathrm{e}^{\mathrm{i}\pi/2},$$

is often cited on this document.

The Dirac $\delta(t)$ distribution can be seen as a supperposition of a continuum of $\cos \omega t$ functions for all frequencies ω. For each fixed frequency (monochromatic) a $\cos \omega t$ has as its integral $\sin \omega t$. We see then that the integral of an even (cosine) function turns into an odd (sine) function. Integration is advancing the phase by $\operatorname{sgn}(\omega)\,\pi/2$. This is an important characteristic of integration filters (as shifting the frequency by $-\operatorname{sgn}(\omega)\,\pi/2$ is a characteristic of derivative filters). To illustrate the concepts about the integration as a digital filter, I will use the FFTLAB software developed by David Hale and distributed with the Seismic Un*x (SU) package. Figure 1.1 shows the integral of a cosine function (a sine function) and its real and imaginary components of its Fourier transform.

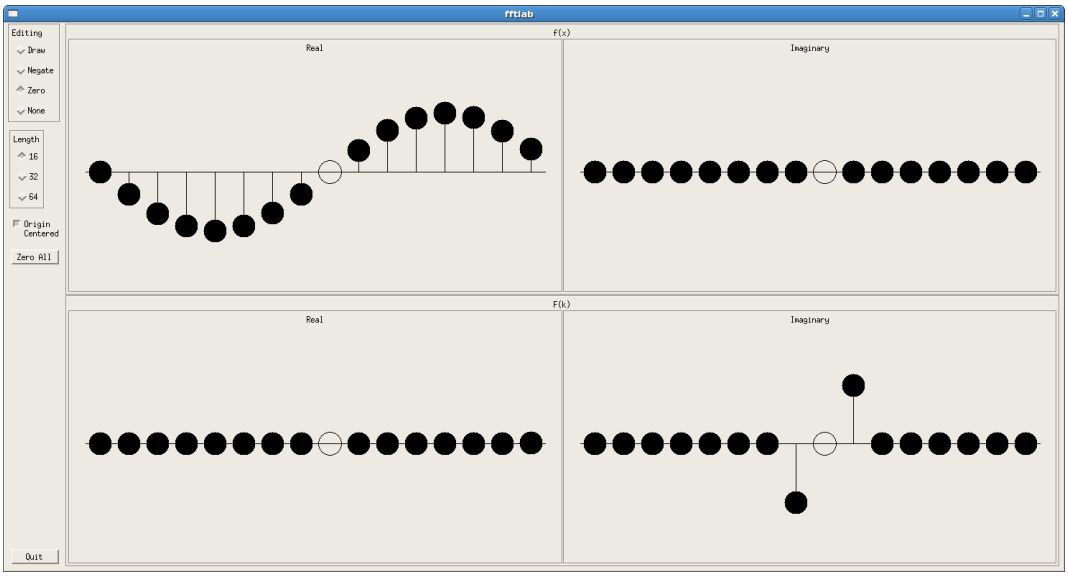


Figure 1.1: The integral of a cosine function and its Fourier transform.

Note that the input signal is real. That is, the imaginary part is zero. In the Fourier domain, the real part is zero and the imaginary part is odd. This is a clear hint that the phase spectrum this filter is $\text{sgn}(\omega)\pi/2$. In general for a digital filter, if the filter is pure real (zero imaginary component), the real component of its Fourier transform is an even function and the imaginary component is odd.

While the Dirac delta could be writen a superposition of a continuum of even functions, for which each of them has an odd integral, the integral of such supperposition is not an odd function, since $H(t)$ is not odd. To understand this loss symmetry we have to redefine the integral operator by having its lower limit at some finite point. Then we redefine 1.1.2 as

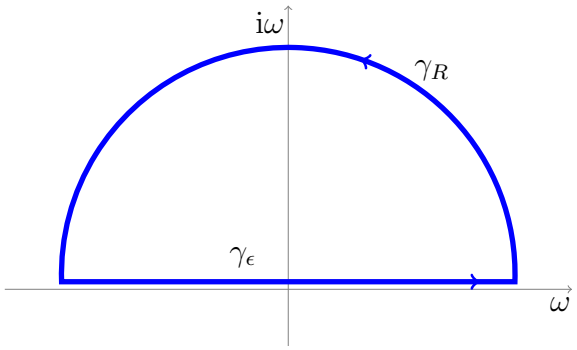
$$F(t) = \int_{t_0}^{t} f(\tau) d\tau.$$

Now, this can be writen as

$$F(t) = \begin{cases} \int_{t_0}^{t} f(\tau) d\tau & \text{for} \quad t > t_0 \\ \\ -\int_{t}^{t_0} f(\tau) d\tau & \text{for} \quad t < t_0 \end{cases}$$

That is, the bilateral integral is splitted in two integrals. A causal integrator (for $t > t_0$) and anti–causal integrator (for $t < t_0$). In this way the function $F(t)$ is an odd function of t with respect to the origin t_0. While symmetry or asymmetry is desired in many situations, all physically realizable filters are causal, and so they can not have either symmetry nor asymmetry.

Figure 1.2: Contour for Fourier transform evaluation (equation 1.1.4) with complex frequency.



This causality behavior is of great importance in dealing with digital signals. We point a few highlights of causal functions. For references about this, the reader is advised to look at Bracewell [3] and Bleistein et. al. [1].

- All digital data is causal. In the sense that there has to be a **first** sample. Data is finite.
- Each causal signal is the product of a function with a Heaviside function.
- If a function is causal, its imaginary Fourier transform component is the Hilbert transform of its real part Fourier transform component.
- If a function is causal, its Fourier transform is analytic in the upper half plane (with our convention of Fourier transform above). The way to understand this problem is by the use of a damping factor (as done in attenuation) by analytically continuing the frequency from the real to the complex plane. Then we assume a new frequency $\Omega = \omega + \mathrm{i}\epsilon$, and consider the integral 1.1.1 with the change this change of variable. That is,

$$f(t) \quad = \quad \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\Omega)\mathrm{e}^{-\mathrm{i}\Omega t}d\Omega. \tag{1.1.4}$$

  Since ϵ is constant, this is a valid statement, but we have to study the new integral as an analytic continuation of the original problem.

  The argument of the exponential function turns out to be

$$-\mathrm{i}\Omega\, t = -\mathrm{i}\omega\, t - \mathrm{i}\ \mathrm{i}\epsilon\, t = -\mathrm{i}\omega\, t + \epsilon\, t$$

  We now consider the contour $\Gamma = \gamma_\epsilon \cup \gamma_R$ indicated in Figure 1.2 and look at the integral as an extended integral on the complex plane of frequencies. We assume that the integration contour passes above all singularities of $F(\omega)$. Since the

integrand is analytic above all it singularities for $t < 0$ then the Cauchy integral evaluates to 0 there. So the $f(t)$ is causal. If there are no singularities we think about this integral in the limit as $\epsilon \to +0$. For $F(\omega)$ then we can pick the contour at $\omega = 0$ and still get the usual Fourier transformation.

The Heaviside function is the signature of causality. Every causal function can be writen as the product of two functions, where one of them is a Heaviside function.

The ideal integrator (a Heaviside function) in the discrete can be represented as a rectangular integrator. This is basically a plain sum.

$$S_k = \sum_{i=0}^{k} a_i \quad ^2$$

for a series $a_0, a_1, \cdots a_n$, Figure 1.3 shows the Fourier response of this integrator (the unit step response).

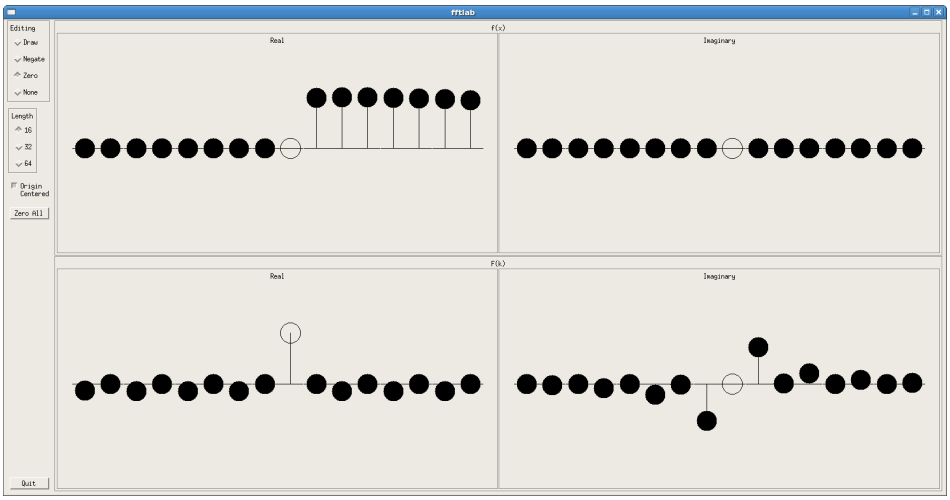


Figure 1.3: The Fourier transform response to a unit step (integrator) function.

The ideal integrator should have a phase of $\text{sgn}(\omega)\pi/2$. as indicated in equation 1.1.3. This means that the real component of its Fourier transform should be zero, each time the imaginary part is non–zero. It is clear from Figure 1.3 that the rectangular filter does not have a clean $\text{sgn}(\omega)\pi/2$ phase behavior, due to small non–zero components of its real Fourier transform. The zeroing of the real components of the Fourier transform produces the bilateral rectangular method filter shown in Figure 1.4.

[2]We use a unit sampling rate for the definitions on this section.

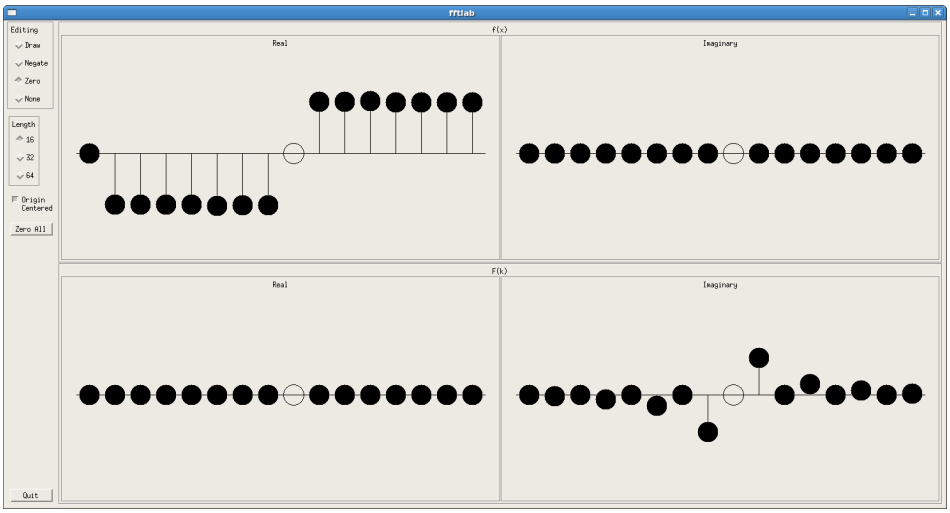


Figure 1.4: The Fourier transform of the bilateral rectangular operator

We will explain later why the rectangular integrator does not have a clean $\pi/2$ phase behavior, proper of integration filters.

The $H(0) = 1/2$ value produces the trapezoidal rule.

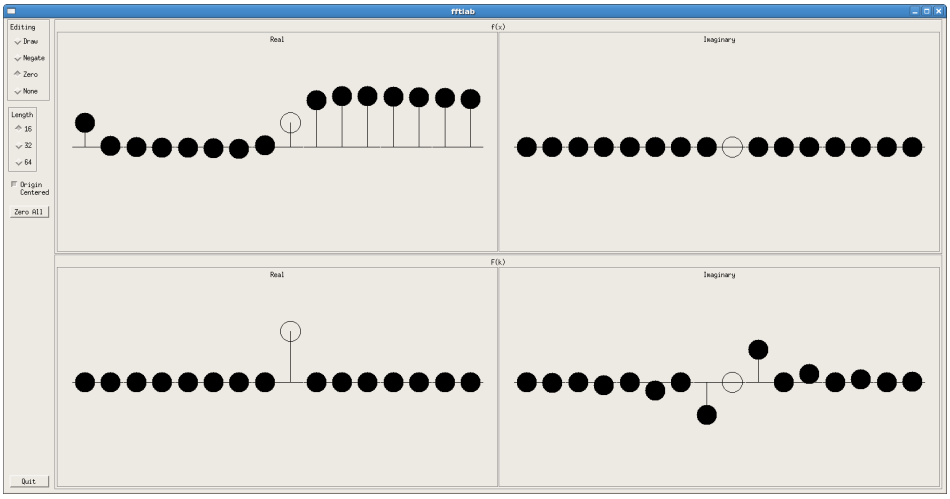


Figure 1.5: The Fourier transform of the trapezoidal rule integrator.

Figure 1.5 shows the impulse response of the trapezoidal rule and its Fourier transform. It is clear here that the $H(0) = 1/2$ fixed the phase issue and provides the ideal integrator spectrum as in equation 1.1.3. Note that the first sample in time domain (the first sample on the left, corresponds to the last live sample if the origin center option is disabled) has the same value as the zero sample which is marked by the empty circle. This is important because the edge effects are stamped on both sides of the finite

data when entering and when leaving the running average sum (another way to see a trapezoidal filter). See equation 1.1.5.

The trapezoidal rule for integration is defined by the equation

$$S_k = \sum_{i=0}^{k-1} \frac{a_i + a_{i+1}}{2} \quad \text{for} \quad k > 0$$

which is a window running average of window size 2, but is is most commonly known for the formula:

$$S_k = \frac{a_0}{2} + \sum_{i=1}^{k} a_i + \frac{a_k}{2}, \tag{1.1.5}$$

So, the trapezoidal rule acknowledges the proper phase behavior of the integrator. However for low frequencies it is unstable. We will search for proper integrators that are causal and stable (that is, minimum phase) and the that is the purpose of the sections ahead.

1.1.1 The relation between analog and digital

The knowledge of a wealth of analog filters in analytical form could be of great use to design digital filters that will produce some desired results. A few highlights which we would like to preserve when converting filters from analog to digital (or vice–versa) are:

- Bandwidth (frequency breadth)
- Amplitude and Phase Spectra.
- Phase behavior. In particular the minimum phase property. This is:
  - Stability.
  - Causality.

In the filter design literature there are several techniques to relate analog to digital filters. Bose, [2] describe the following techniques to convert between analog and digital filters:

- Use the bilinear transformation
- Impulse Invariant Method
- Numerical Integration Method
- Matched Z–transforms.

Bose (in his Table 3.6) presents advantages and disadvantages of each of these methods in terms of the highlights itemized above and in addition to difficulties on implementation and aliasing contamination. I believe that the best approach for numerical integration is the bilinear transformation, that according to some authors (see for example, Oppenheim et. al., [5]) is motivated by the trapezoidal rule of integration. [3]

The transformation from analog to digital is done by mapping the Laplace parameter $s = \mathrm{i}\omega$, on the continuum to the Z–transform parameter $z = \mathrm{e}^{s\Delta t}$ on the discrete, with sampling rate Δt.

The digital representation of filters is through the Z–transform. Polynomial representations are easy to implement, but many filters come in fractions of polynomials. Filters writen in terms of fractions of polynomials are evaluated recursively. These are the type of filters used for integration (accumulators). The simplest rational function of polynomials is a bilinear representation (bilinear here means that the numerator and denominator are linear functions of the z argument. The bilinear function is not even linear).

The idea is to find a rational approximation for the parameter s as a function of z. Let us start by finding first a rational function of z as a function of s (Claerbout, [4]).

$$\begin{aligned} z &= \mathrm{e}^{s\Delta t} \\ &= \frac{\mathrm{e}^{s\Delta t/2}}{\mathrm{e}^{-s\Delta t/2}} \\ &\approx \frac{1 + s\Delta t/2}{1 - s\Delta t/2}. \end{aligned}$$

Then from here

$$z(1 - s\Delta t/2) \approx 1 + s\Delta t/2,$$

that is

$$1 - z \approx -\frac{s\Delta t}{2}(1 + z)$$

and

$$s \approx -\frac{2}{\Delta t}\frac{1 - z}{1 + z}.$$

[3] In fact, I will go the opposite way. I derive the bilinear approximation and then show that it represents the difference equation for the trapezoidal rule of numerical integration.

This approximation of s as the ratio of two linear z transforms is called the bi–linear transformation. Let us assume $z = a + \mathrm{i}b$, with $a, b \in \mathbb{R}$. Then,

$$\begin{aligned} s &\approx -\frac{2}{\Delta t}\frac{(1-a)-\mathrm{i}b}{(1+a)+\mathrm{i}b} \\ &= -\frac{2}{\Delta t}\frac{[(1-a)-\mathrm{i}b][(1+a)-\mathrm{i}b]}{(1+a)^2+b^2} \\ &= -\frac{2}{\Delta t\,[(1+a)^2+b^2]}[(1-a^2-b^2-\mathrm{i}b(1+a+1-a)] \\ &= -\frac{2}{\Delta t\,[(1+a)^2+b^2]}(1-a^2-b^2-2\mathrm{i}b) \\ &= -\frac{2(1-a^2-b^2)}{\Delta t\,[(1+a)^2+b^2]} + \mathrm{i}\,\frac{4b}{\Delta t\,[(1+a)^2+b^2]} \end{aligned}$$

The amplitude is approximated by

$$|s| \approx \frac{2\sqrt{(1-a^2-b^2)^2+4b^2}}{\Delta t[(1+a)^2+b^2]} = \frac{2\sqrt{1-2a^2+2b^2+2a^2b^2+a^4+b^4}}{\Delta t[(1+a)^2+b^2]}$$

and the phase is approximated by

$$\phi(s) \approx \arctan\left(\frac{2b}{a^2+b^2-1}\right). \tag{1.1.6}$$

If $z = \exp(\mathrm{i}\omega\Delta t)$, then $a = \cos\omega\Delta t$ and $b = \sin\omega\Delta t$, and $1 - a^2 - b^2 = 0$, then

$$s \approx \mathrm{i}\frac{4b}{\Delta t[(1+a)^2+b^2]}$$

and with

$$(1+a)^2 + b^2 = 2\cos\omega\Delta t, \qquad b = \sin\omega\Delta t$$

then

$$s \approx \frac{2}{\Delta t}\,\mathrm{i}\tan(\omega\Delta t/2). \tag{1.1.7}$$

from which its amplitude and phase responses are given by

$$|Y(f)| = \frac{2}{\Delta t}|\tan(\pi f\Delta t)| \qquad \phi(f) = \mathrm{sgn}\omega\,\pi/2.$$

A geometrical way to understand the amplitude and phase behavior is by thinking of z values as two–dimensional vectors. Figure 1.6 shows a point z and the vectors $z + 1$ (red) and $z - 1$ (green). The numerator is represented by $z - 1$ as the green line, the denominator $z + 1$ is the red line. The amplitude of the filter is dominated

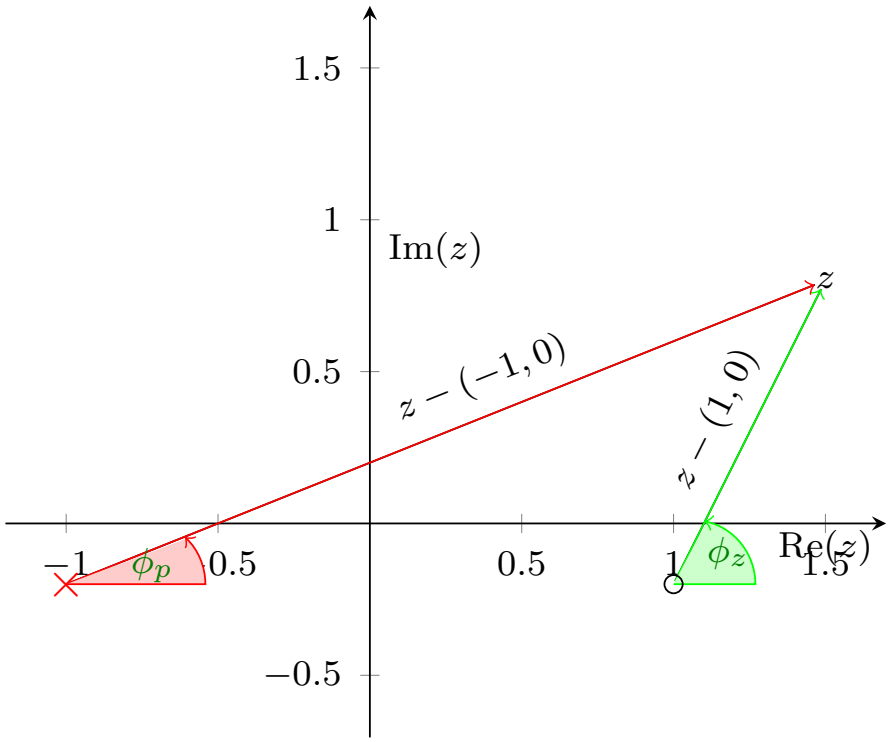


Figure 1.6: This figure aids in the geometrical illustration of the phase and amplitude spectra from associating z to a vector. The zero is represented by a small circle and the pole by a "$\times$" sign. Here we observe the ratio $(z-1)/(z+1)$. The amplitude is the ratio of the norms of the green over the red vector. The phase is the difference between the green (ϕ_z) and the red (ϕ_p) phases.

by the pole. That is, as z gets close to the pole $(-1, 0)$ the amplitude grows without bounds. Of course for z close to the zero, the amplitude should be low, but we are not much interested on the amplitude spectrum at this moment. We will come to this later with an analytic approach. The phase is the difference between the phase of these two vectors. At any point z in the upper half plane , the phase is the difference

$$\phi = \phi_z - \phi_p,$$

and at any point z in the lower half plane, the phase is the difference

$$\phi = -\phi_z + \phi_p.$$

Since the upper half plane is associated with $\omega > 0$ and the lower half plane with $\omega < 0$ we can write compactly

$$\phi = \operatorname{sgn}\omega(\phi_z - \phi_p).$$

Along the real line inside the unit circle and coming from above (that is for $-1 \leq \mathrm{Re}(z) \leq 1, \mathrm{Im}(z) > 0$), $\phi_z = \pi$ and $\phi_p = 0$, so $\phi = -\pi$. Now if $\mathrm{Im}(z) < 0$ then $\phi = -\pi$. This means that the segment $-1 \leq z \leq 1$ is a branch cut. On the real line but on the right of the zero $z = (1, 0)$, both angles ϕ_z and $\phi_p = 0$, so $\phi = 0$. On the real line, to the left of the pole $z = (-1, 0)$ both angles are π so also $\phi = 0$.

Let us now study the phase along the unit circumference. From elementary geometry, the angle at z is always $\pi/2$ and

$$\phi_z = \phi_p + \pi/2,$$

so

$$\phi_z - \phi_p = \pi/2.$$

That is, along the unit circumference the phase is $\operatorname{sgn}\omega\,\pi/2$. Figure 1.7 illustrates this situation.

Digital filters can be easily explained based on their zero–pole plots. The zero–pole plot for the digital integration filter has a zero at -1 and a pole at 1, as shown in Figure 1.8 .

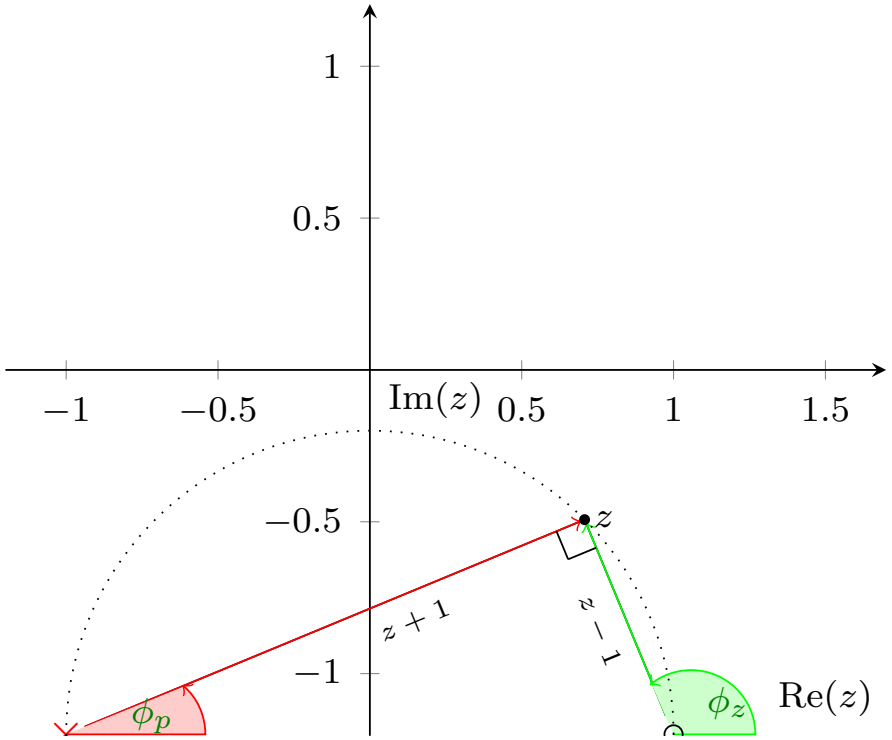


Figure 1.7: Phase of bilinear transformation $(z-1)/(z+1)$ with $z = \exp(\mathrm{i}\omega\Delta t)$. Here $1 = (1, 0)$. Clearly along the upper semi–circumference $\phi = \phi_z - \phi_p = \pi/2$.

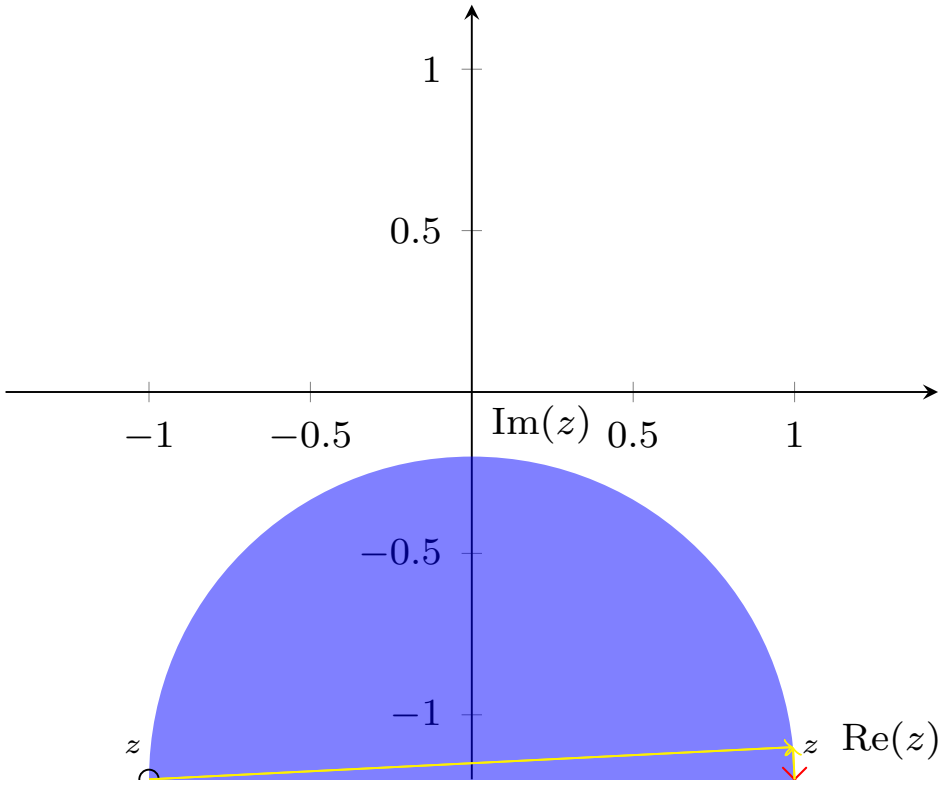


Figure 1.8: Low frequency enhance integrator filter. Here $z = \mathrm{e}^{\mathrm{i}\omega\Delta t}$. The zero is represented by a small circle and the pole by a "×" sign.

Here we use phase plots to illustrate the filters responses. Phase plots come in many flavors. Wegert and Semmler, [7] [4] show the power and art of phase plots to illustrate complicated complex functions. Poles, zeros, brunch cuts, essencial singularities, etc. All can be easily identified with the help of phase plots. To understand the coloring system, I show in figures 1.9 and 1.10 the phase plots for the fucntions z (left) and $1/z$ (right). In all cases the C function `-atan2` was used. We picked three color plots: red, green and blue and color bars displaying the phases between $-\pi$ and π.

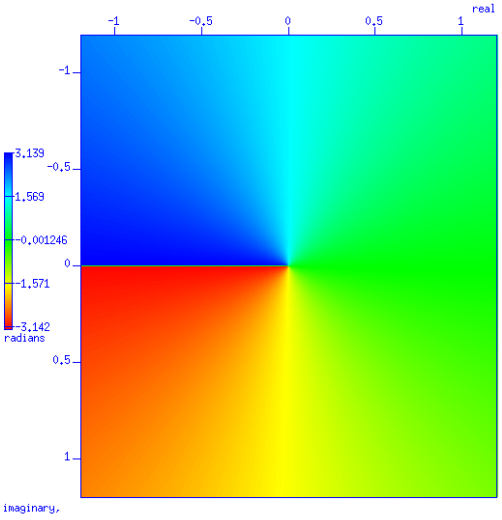


Figure 1.9: Phase plot for the function $f(z) = z$. Angle increases in the counter clockwise direction.

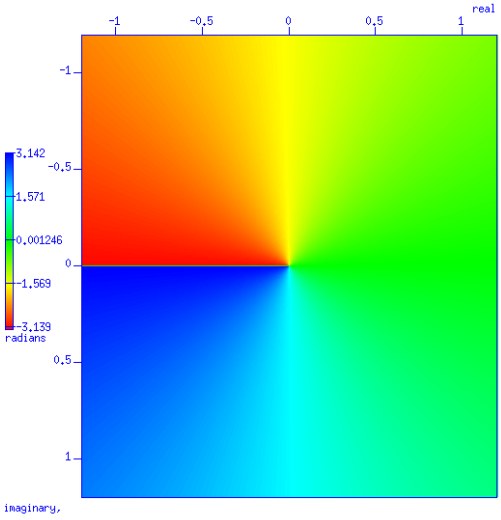


Figure 1.10: Phase plot for the function $f(z) = 1/z$. Angle decreases in the counter clockwise direction.

Close to zero, the identity map $f(z) = z$ has all colors between green (low) to blue (high positive) to red as we wind around zero in a counter clockwise direction. Close to zero the inverse function $f(z) = 1/z$ has a pole and all the colors between green to red (high negative) to blue. Which is what we expect since for $z = A \exp^{\mathrm{i}} \phi)$

$$\arg(z) = \phi \qquad \arg(1/z) = -\phi.$$

A lower resolution plot where we can see the color divisions shows the mapping of constant phase lines which helps even more to understand the phase flow. Figures 1.11 and 1.12 show the phase plots for $f(z) = z$ and $f(z) = 1/z$, this time in a lower resolution scale that let us see the phase transitions, which in these two functions are radias but as we will obverse later (see for example Figure 1.14) could be more paths.

[4] http://www.arxiv.org/pdf/1007.2295

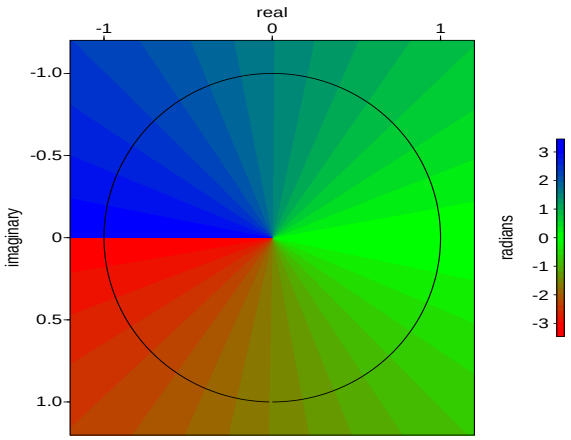


Figure 1.11: Phase plot for the function $f(z) = z$. Angle increases in the counter clockwise direction. Lines of constant phase are seen as radial lines.

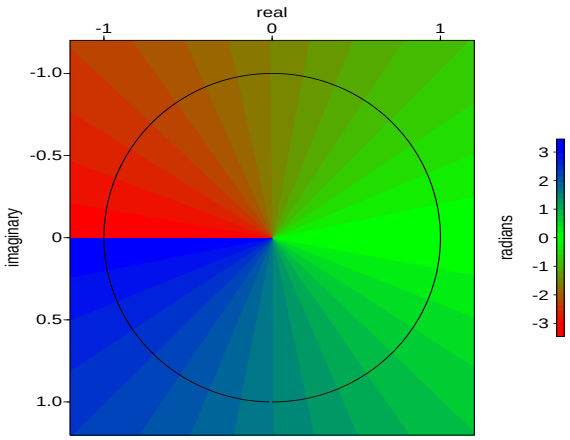


Figure 1.12: Phase plot for the function $f(z) = z$. Angle decreases in the counter clockwise direction. Lines of constant phase are seen as radial lines.

Note that the unit circumference is included since it will be the reference contour for the Z transform. Figure 1.13 shows the phase function 1.1.6.

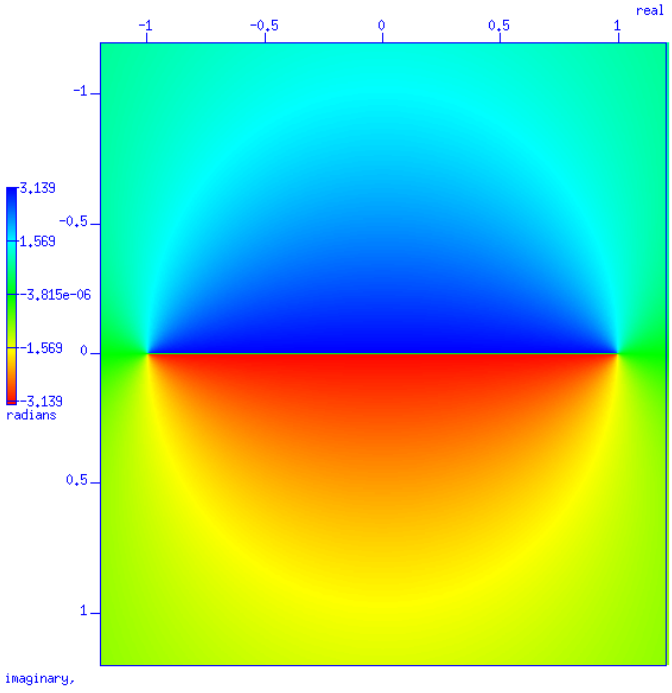


Figure 1.13: Phase plot of the bilinear transformation.

Phase plots are good because they highlight important features of their corresponding complex mappings. Referring to the plot in Figure 1.13 we observe:

- The zero at $z = 1$ and the pole at $z = -1$. Note that the sequence of phases as we rotate counter–clockwise around the zero is reversed as compared with the same oriented rotation around the pole. There is not a clear definition of the phase at a zero and a pole. All phases take place around them. What is the phase of a zero vector? Had we chosen the electrical engineers convention for z^{-1} instead of z, those zeroes and poles inside the unit circle would be mapped outside to their reciprocals, and those outside to the inside.

- The line between the the pole $(-1, 0)$ and the zero $(1, 0)$ is a brunch cut of the bilinear function.

- From equation 1.1.6 we observe that the phase is even (symmetric) with respect to the real axis (a) and odd with respect to its imaginary axis (b) as shown in the picture.

- The phase is zero in the real axis. At a small $b > 0$, the phase is close to π inside the unit circle (where $a^2 + b^2 - 1 < 0$), and for small $b < 0$, the phase is close to $-\pi$ inside the unit circle. Outside of the unit circle the values of the phase are close to 0 above (for $b > 0$) and zero below (for $b < 0$). This is the case of the regular $\arctan(y/x)$ evaluation.

- At the unit circle ($z = \exp(\mathrm{i}\omega\Delta t)$), the phase is constant $\pi/2$ for $b > 0$ and $-\pi/2$ for $b < 0$. This means that the real component of s is zero and that the imaginary axis in the s plane is mapped into the unit circle in the z plane.

- If we take a point on the phase plot for which the phase is $0 < \phi(s) < \pi/2$ that point is outside of the unit circle, and since, in the s plane a point with phase $0 < \phi(s) < \pi/2$ is located in the right of s plane, we conclude that the outside of the unit circle on the z plane maps to the right side of the s plane. Likewise, by using a similar argument, the inside of the unit circle maps into the left side of the real plane. This is very important since it indicates that the bilinear transformation preserves the minimum phase filter condition (all zeros and poles outside of the unit circle in the z plane and on the right of the s plane).

In addition to the properties of the bilinear transformation listed above, this method of filter design is immune to aliasing. However, we should be aware of the fact that the frequencies are mapped non–linearly. From equation 1.1.7 we find that

$$s = \mathrm{i}\omega_a \approx \frac{2}{\Delta t}\,\mathrm{i}\tan(\omega\Delta t/2). \tag{1.1.8}$$

where ω_a is the frequency corresponding to the analogue filter. We find the relation (and its inverse) between the analog and digital frequencies:

$$\begin{aligned} \omega_a &= \frac{2}{\Delta t}\tan\left(\frac{\omega\Delta t}{2}\right) \\ \omega &= \frac{2}{\Delta t}\arctan\left(\frac{\omega_a\Delta t}{2}\right) \end{aligned}$$

An important observation is that the variation of analog frequencies between $-\infty < \omega_a < +\infty$ will map the discrete range of frequencies $-\pi/2 < \omega < \pi/2$, with Nyquist extrema. This frequency compression (or stretching if looked in the other direction) is known as *frequency warping* .

The design of a digital filter with transfer function $Y_d(\omega_d)$ from an analog filter with transfer function $Y_a(\omega_a)$ is achieved with the substitution

$$Y_d(\omega) = Y_a(s)\Big|_{s=-\frac{2}{\Delta t}\frac{1-z}{1+z}}$$

with $z = \exp(\mathrm{i}\omega\Delta t)$.

Returning to the integration theory, we could observe two basic types of numerical integrations. The analog integration and the digital integration. While physical filters are analog by nature, the word "analog" here means that the computer implementation is a direct discretization of the continuous Laplace operator $1/\mathrm{i}\omega$ as shown in the next section. The Laplace transform signature $s = \mathrm{i}\omega$ in the continuous domain is mapped to the Z–transform signature $z = \exp(\mathrm{i}\omega\Delta t)$ in the discrete domain with the sampling rate Δt. In the following sections we show the specifics of integration under each category.

1.2 Analog

We know that the Laplace parameter s has the following mapping into the time domain

$$Y(s) = s = -\mathrm{i}\omega \overset{\mathcal{L}^{-1}}{\rightarrow} \frac{d}{dt}$$

so, since integration is the inverse of differentiation, no wonder

$$Y(s) = \frac{1}{s} = -\frac{1}{\mathrm{i}\omega} \overset{\mathcal{L}^{-1}}{\rightarrow} \int^t$$

This equation shows that the analog filter can be applied in the frequency domain by dividing by $-s = -\mathrm{i}\omega$. Of course, these needs a tapering window (Hamming, Kaiser or any other), to smooth out edge effects. Note that the function as a pole at $\omega = 0$. That is it amplifies the zero frequency without bounds and it is a low frequency amplifier. Integration are smoothing filters, that focus on the low frequencies, while derivations are sharping filters that amplify high frequencies.

1.3 Digital

The transformation of an analog filter to a digital can be done through the bilinear mapping

$$s \mapsto -\frac{2}{\Delta t}\frac{1-z}{1+z}, \tag{1.3.9}$$

with $z = \mathrm{e}^{\mathrm{i}\,\omega\Delta t}$.

For the integration filter we have

$$Y_a(s) = -\frac{1}{s} \approx \frac{\Delta t}{2}\frac{1+z}{1-z} = Y_d(z) \tag{1.3.10}$$

However, as I illustrated above, the phase plot (Figure 1.13) conveys much more information that a zero–pole plot. So in what follows I will use phase plots to explain filter attributes.

We show next that this filter corresponds to the well known trapezoidal integration method.

The recursive implementation: From the transfer function definition $Y(z) = \frac{\mathrm{P}_{out}(z)}{\mathrm{P}_{in}(z)}$ we find

$$\frac{\mathrm{P}_{out}(z)}{\mathrm{P}_{in}(z)} = \frac{\Delta t}{2}\frac{1+z}{1-z}$$

and from here (For clarity, I omit the argument z in P_{in} and P_{out})

$$\mathrm{P}_{out}(z) - z\mathrm{P}_{out}(z) = \frac{\Delta t}{2}\mathrm{P}_{in}(z) + z\frac{\Delta t}{2}\mathrm{P}_{in}(z)$$

and so the implementation algorithm is

$$\mathrm{P}_{out}[i] = \mathrm{P}_{out}[i-1] + \frac{\Delta t}{2}\left(\mathrm{P}_{in}[i] + \mathrm{P}_{in}[i-1]\right) \tag{1.3.11}$$

which is precisely the trapezoidal rule of integration. Figure 1.5 shows the impulse response of the trapezoidal rule for integration, as well as its Fourier response.

We note that this corresponds to the causal and anticausal integrator with 1/2 in its zero and last samples.

The Simpson's rule [5] is based on the evaluation

$$\int_a^{a+2\Delta t} f(t)dt \approx \frac{\Delta t}{3}[f(a) + 4f(a+\Delta t) + f(a+2\Delta t).$$

[5] http://ezekiel.vancouver.wsu.edu/~ cs330/lectures/integration/simpsons.pdf

The derivation of this rule is obtained by fitting a quadratic polynomial into the function and finding its coefficients A, B, C such that the polynomial matches the function at a, $a + \Delta t$ and $a + 2\Delta t$.

A recursive implementation of the Simpson's rule is based on the equation

$$y_{n+2} = y_n + \frac{\Delta t}{3}[x_n + 4x_{n+1} + x_{n+2}],$$

from which the transfer function for the Simpson's rule is given by

$$Y_s(z) = \frac{\Delta t}{3}\frac{1 + 4z + z^2}{z^2 - 1} \tag{1.3.12}$$

Tompkins, [6] website [6] claims that the Simpson's rule is the most common form of numerical integration). Figure 1.14 shows a phase plot of the Simpson rule.

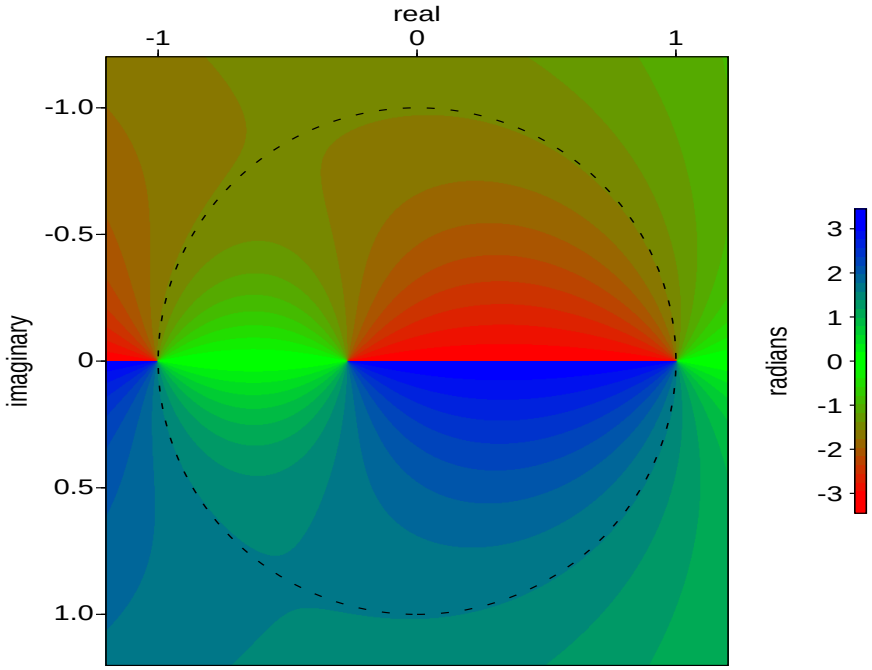


Figure 1.14: Phase plot of the Simpson's rule transfer function

The zeroes of the transfer function 1.3.12 are given by

$$1 + 4z + z^2 = 0 \quad \Rightarrow \begin{cases} z = -2 - \sqrt{3} \approx -3.7331 \\ z = -2 + \sqrt{3} \approx -0.2675 \end{cases}$$

[6] http://www.scribd.com/doc/79208926/Bio-Medical-Signal-Processing-Tompkins

and the poles by $z^2 - 1 = 0, \Rightarrow z = \pm 1$. The zero $-2 - \sqrt{3}$ is not shown in the figure. We can clearly see the zero $z = -2 + \sqrt{3}$ and the two poles. The phase along the unit circle is sgn$\omega\, \pi/2$. We proof this in section 1.4.2. In the phase plots, the circular frequency ω has the range

$$-\pi < \omega \Delta t \leq \pi$$

with π being the normalized to radians Nyquist frequency (the point $(-1, 0)$). The positive frequencies are in the upper semi–circle while the negative frequencies are in the lower semi–circle. Note that the zeroes wind clock–wise while the poles wind counter–clockwise. This is the opposite winding direction to what we expected. Why? Clearly the Simpson's integration filter is not a minimum phase filter. The two zeroes on top ot the unit circumference and another zero inside the unit circle, destroy stability. Hence, in spite that the Simpson's integration preserves the correct phase of sgn$\omega\ \pi/2$ is not stable and is not recommended for integrating signals.

As a final method, I show the leaky integration. In his Stanford Exploration Project SEP [7] website, Jon Claerbout introduces an algorithm that he calls *leaky integration* .

The idea is simply as follows. A two point backward differentiator operator in the Z domain is written as

$$D(z) = 1 - z$$

the inverse of this operator (an integration) is then

$$I(z) = \frac{1}{1 - z} = 1 + z + z^2 + \cdots \tag{1.3.13}$$

which is the causal integrator (a discrete Heaviside function). This is the causal rectangular integration. As done above, we can find a parameter ρ to move the zero outside the unit circle.

$$I_{leaky}(z) = \frac{1}{1 - \rho\, z} = 1 + \rho z + \rho^2 z^2 + \cdots \tag{1.3.14}$$

What happens in the leaky integration algorithm is that the zero is removed and so the filter was forced into a minimum phase filter which is both causal and stable.

Figure 1.15 shows the impulse response of the leaky integrator for $\rho = 0.9$

[7] http://sepwww.stanford.edu/sep/prof/pvi/zp/paper_html/node2.html

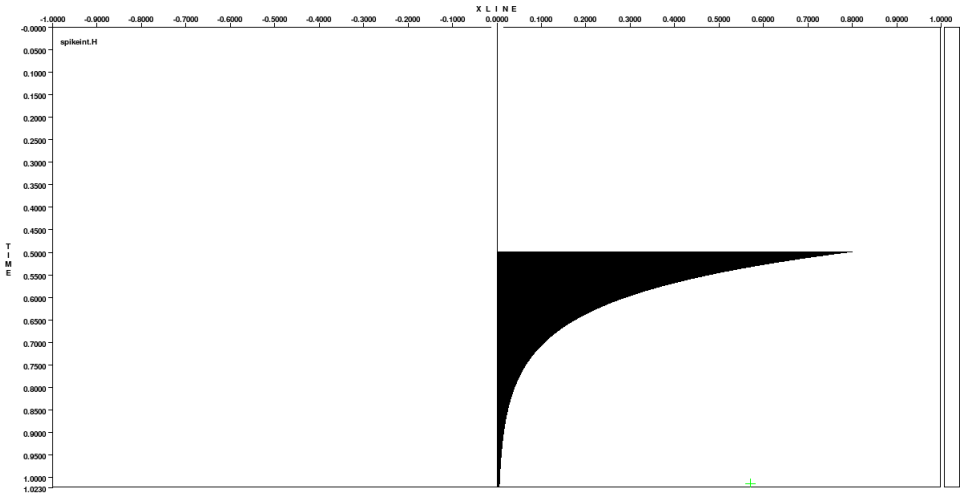


Figure 1.15: Impulse response for the leaky integration ($\rho = 0.9$). Note that it is causal ($t_0 = 0.5$s) and stable. That is, it is a minimum phase filter.

where the zero time is set to $t_0 = 0.5s$. We note that the response is both, causal and stable. The filter is minimum phase.

1.4 The Spectrum

1.4.1 Analog

From the analog transfer function

$$Y(\omega) = -\frac{1}{\mathrm{i}\omega}$$

it is clear that the integration filter has a $\operatorname{sgn}\omega$ $\pi/2$ phase (each $\cos\omega\Delta t$) turns into a $\sin\omega\Delta t$. This is a constant phase retardation of $\pi/2$ for each frequency component. The amplitude spectrum is obviously given by

$$|Y_a(\omega)| = \frac{1}{\omega}$$

for which the low frequency components diverge as we approach $\omega = 0$, producing an infinite DC bias.

To calibrate the filter I pick a frequency of 10 Hz. That is

$$Y_a(2\pi f)\Big|_{f=10Hz} = \frac{1}{62.83185307179586476880} = .01591549430918953357$$

so

$$A_a = 20\log_{10} Y_a(2\pi f) \approx -35.963597367162300995 00. \tag{1.4.15}$$

1.4.2 Digital

The leaky integration formula 1.3.14 is missing the Δt factor that comes naturally as a byproduct of the discretization. This factor is important to preserve the correct scaling of the filter. So, I write instead

$$Y_{leaky}(z) = \frac{\Delta t}{1 - \rho z},$$

and so

$$|Y_{leaky}(f)| = \left| \frac{\Delta t}{1 - \rho z} \right| = \left| \frac{1}{1 - \rho \mathrm{e}^{\mathrm{i}2\pi f \Delta t}} \right| = \frac{\Delta t}{\sqrt{1 + \rho^2 - 2\rho \cos 2\pi f \Delta t}}.$$

Then

$$A_{cal} = 20 \log_{10} Y_d(f) = -60 - 10 \log_{10}(1 + \rho^2 - 2\rho \cos \omega \Delta t)$$

and at $f = 10$ Hz

$$A_{cal} \approx -35.962168548578000745 90$$

which compares well to formula 1.4.15 up to 4 digits.

To find the phase of the spectrum, we write

$$I_{leaky}(z) = \frac{1}{1 - \rho \cos \omega \Delta t - \rho \mathrm{i} \sin \omega \Delta t} = \frac{1 - \rho \cos \omega \Delta t + \rho \mathrm{i} \sin \omega \Delta t}{(1 - \rho \cos \omega \Delta t)^2 + \rho^2 \sin^2 \omega \Delta t}.$$

That is,

$$I_{leaky}(z) = \frac{1 - \rho \cos \omega \Delta t}{1 - \rho^2 - 2\rho \cos \omega \Delta t} + \mathrm{i} \frac{\rho \sin \omega \Delta t}{1 - \rho^2 - 2\rho \cos \omega \Delta t}$$

and so the phase spectrum is given by

$$\phi_{leaky}(f) = \arctan \left(\frac{\rho \sin 2\pi f \Delta t}{1 - \rho \cos 2\pi f \Delta t} \right)$$

Note that due to the real component contribution, the phase departs from the ideal $\mathrm{sgn}(\omega)\pi/2$ except for the zero frequency (and $\rho = 1$). Figure 1.15 shows the impulse response for a leaky integrator (the zero time $t_0 = 0.5$ to show his causal behavior). Note that it is causal and stable (energy converges as $t \to \infty$).

The corresponding phase plot for the leaky integrator transfer function with $\rho = 0.9$ is shown in figure 1.16.

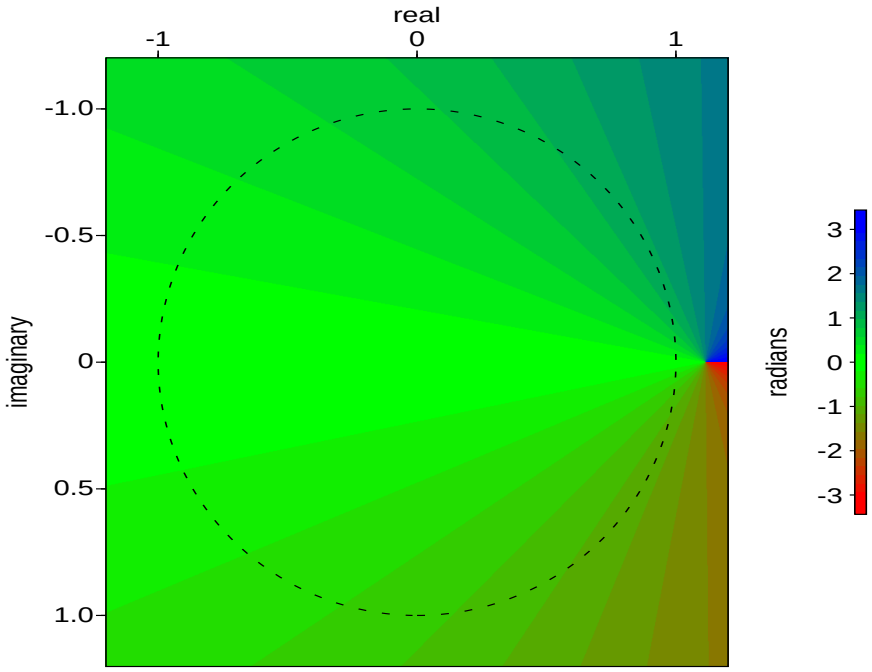


Figure 1.16: Phase plot for the leaky integration transfer function ($\rho = 0.9$). Note that the pole is pulled out of the unit circle, making of this a minimum phase filter (causal and stable). However the phase on each point of the circle is in size smaller than $\pi/2$.

For $\rho = 1$ the leaky integration is equivalent to rectangular algorithm, so the amplitude spectrum for the rectangular algorithm is

$$A_r(f) = \left| \frac{\Delta t}{\sqrt{2 - 2\cos 2\pi f \Delta t}} \right| = \left| \frac{\Delta t}{2 \sin \pi f \Delta t} \right| = \frac{\Delta t}{2} |\csc \pi f \Delta t 2|,$$

while the phase spectrum is given by

$$\tan \phi_r(f) = \frac{\sin 2\pi f \Delta t}{1 - \cos 2\pi f \Delta t} = \frac{\sin 2\pi f \Delta t}{2 \sin^2(\pi f \Delta t)} = \frac{2 \sin(\pi f \Delta t) \cos(\pi f \Delta t)}{2 \sin^2(\pi f \Delta t)} = \cot \pi f \Delta t$$

and since

$$\cot x = \tan(\pi/2 - x),$$

$$\phi_r(f) = \frac{\pi}{2} - \pi f \Delta t$$

So the rectangular algorithm does not acknowledge the right phase behavior except at the zero frequency, and departs from it linearly. At 0 frequency (the point $(1, 0)$) the phase is the desired $\pi/2$ but as the frequency increases toward Nyquist $f = 1/(2\Delta t)$, the phase decays to 0. Figure 1.3 shows the non–zero phase components from the Fourier real and imaginary components. Figure 1.17 shows the phase plot for the rectangular integration filter. It is shown in the plot how the phase decays from $\mathrm{sgn}\omega\,\pi/2$ toward 0. We observe also the pole at the point $z = 1$. The correct interpretation is that the phase goes from $-\pi$ to π as we rotate around the point in counter–clockwise direction.

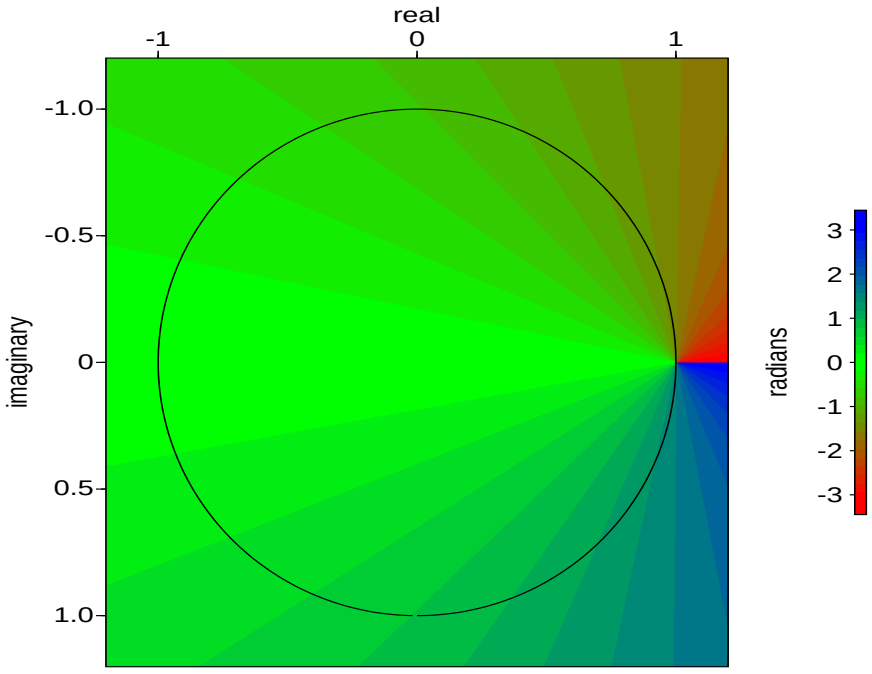


Figure 1.17: Phase plot for the rectangular integrator.

We observe that close to the pole the phase is $\mathrm{sgn}\omega\pi/2$ and that it starts decreasing, when moving along the unit circle towards the Nyquist (the point $(-1, 0)$) where

$$\phi_r(|f_N|) = \pi \mathrm{sgn}(f) \frac{1}{2\Delta t} \Delta t - (\mathrm{sgn} f)\ \frac{\pi}{2} = 0.$$

For the trapezoidal; from equation 1.1.8 after inversion we find

$$-\frac{1}{s} \approx \frac{\mathrm{i}\,\Delta t}{2} \cot \frac{\omega \Delta t}{2}.$$

Note that the low frequency approximating $\cos(\omega\Delta t/2) \approx 1$ and $\sin(\omega\Delta t/2) \approx \omega\Delta t/2$, yields back $-1/(\mathrm{i}\omega)$, that is $-1/s$.

The amplitude and phase spectrum are given by

$$A_t(\omega) = \frac{\Delta t}{2}\left|\cot\frac{\omega\Delta t}{2}\right| \qquad \phi_t(\omega) = \mathrm{sgn}(\omega)\frac{\pi}{2}.$$

Finally for the Simpson's rule. We find the amplitude and phase spectrum of the transfer function 1.3.12. Expanding for $z = \exp \mathrm{i}\omega\Delta t$,

$$Y_S(f) = \frac{\Delta t}{3}\frac{1 + 4\cos\omega\Delta t + \cos 2\omega\Delta t + \mathrm{i}(4\sin\omega\Delta t + \sin 2\omega\Delta t)}{-1 + \cos 2\omega\Delta t + \mathrm{i}\sin 2\omega\Delta t} \tag{1.4.16}$$

From

$$\begin{aligned} 1 + \cos 2x &= 1 + \cos^2 x - \sin^2 x = 2\cos^2 x \\ 1 - \cos 2x &= 1 - \cos^2 x + \sin^2 x = 2\sin^2 x \end{aligned}$$

we simplify 1.4.16 to

$$\begin{aligned} Y_S(f) &= \frac{\Delta t}{3}\frac{2\cos^2\omega\Delta t + 4\cos\omega\Delta t + 2\mathrm{i}(2\sin\omega\Delta t + \sin\omega\Delta t\cos\omega\Delta t)}{-2(\sin^2\omega\Delta t - \mathrm{i}\sin\omega\Delta t\cos\omega\Delta t)} \\ &= \frac{-\Delta t}{3\sin\omega\Delta t}\frac{\cos^2\omega\Delta t + 2\cos\omega\Delta t + \mathrm{i}\sin\omega\Delta t(2 + \cos\omega\Delta t)}{\sin\omega\Delta t - \mathrm{i}\cos\omega\Delta t} \\ &= \frac{-\Delta t}{3\sin\omega\Delta t}\frac{\cos\omega\Delta t(\cos\omega\Delta t + 2) + \mathrm{i}\sin\omega\Delta t(2 + \cos\omega\Delta t)}{\sin\omega\Delta t - \mathrm{i}\cos\omega\Delta t} \\ &= \frac{-\Delta t(\cos\omega\Delta t + 2)}{3\sin\omega\Delta t}\frac{\cos\omega\Delta t + \mathrm{i}\sin\omega\Delta t}{\sin\omega\Delta t - \mathrm{i}\cos\omega\Delta t} \\ &= -\mathrm{i}\frac{\Delta t(\cos\omega\Delta t + 2)}{3\sin\omega\Delta t} \end{aligned}$$

So the amplitude and phase spectra of the Simpson's rule is given by

$$|Y_S(f)| = \left|\frac{\Delta t[2 + \cos(2\pi f\Delta t)]}{3\sin 2\pi f\Delta t}\right| \qquad \phi_s(f) = -\mathrm{sgn}\,\omega\,\frac{\pi}{2}.$$

Then the Simpson's rule yields a phase of opposite sign as the ideal integration should provide.

Figure 1.18 shows the amplitude spectra of the analog and the digital integration filters explained in this document. The digital filters are the leaky integration (blue), the rectangular (purple), the trapezoidal (light blue) and the Simpson (orange). Also the calibration point of $10Hz$ is shown for reference.

Notes on Integration for Digital Data

Herman Jaramillo
GXT ION—
2105 Citywest Blv # 400, Houston, TX 77042

August 6, 2013

Figure 1.18: Amplitude Response of analog ideal (red) and digital: leaky in blue, rectangular in purple, trapezoidal in light blue and Simpson in orange; integration filters.

We observe that in the seismic range between about 2 Hz up to about 100 Hz, they all have a consistent behavior. However the leaky integrator is the only integrator that is stable for low frequencies, while all others scale the low frequencies without bounds. For high frequencies (above 100 Hz) the Simpson, rule is the most unstable. It can amplify noise above 100 Hz. The rectangular filter seems the most adequate to avoid amplification of high frequency noise. It is perhaps more an art than a science to pick the right filter with the right parameters and this corresponds to the field of processing. I provide these plots as an aid to take a decision or to design Quality Control (QC) tools to better handle digital data. As a final example, Figure 1.19 show the spectra of the ideal analog versus two main digital: The trapezoidal and the leaky integration filters. This time the spectra is computed directly on the digital impulse responses computed from a C code and written in a file.

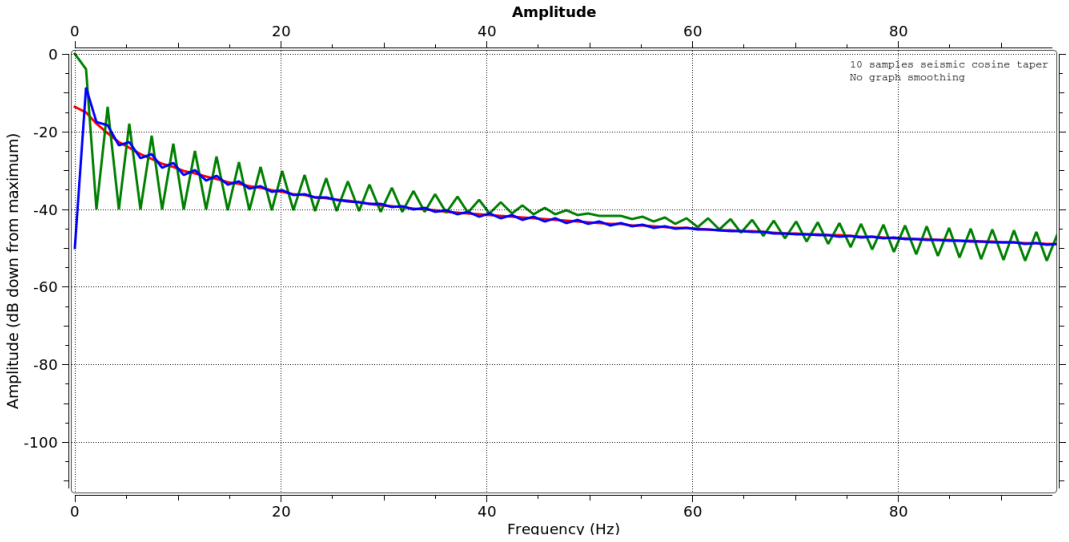


Figure 1.19: Red is the analogue filter with amplitude $1/\omega$. The green is the trapezoidal filter and the blue is the leaky filter.

We observe that the leaky integration filter is smoother in terms of oscillations and desirable.

Next I show the implementation algorithms for the leaky and trapezoidal integration filters.

1.5 Implementation

From $Y(z) = \frac{\mathrm{P}_{out}(z)}{\mathrm{P}_{in}(z)}$ we find (omitting the argument "z")

$$\mathrm{P}_{out}(1 - \rho z) = \mathrm{P}_{in}\Delta t$$

and then the algorithm is given by

$$\mathrm{P}_{out}[i] = \rho \mathrm{P}_{out}[i-1] + \mathrm{P}_{in}[i]\,\Delta t$$

The C code implementation for the trapezoidal and the leaky integration is below

```
/* trapezoidal */
float  * intfilt2( float *input, int nt, float sampfrq, float rho)
{
    float *aprime;
    int ii;
    double dt = 0.5/sampfrq;
```

```
    aprime = ealloc1float(nt);
    memset((void *) aprime, 0 , nt*FSIZE);


    for (ii=1; ii<nt; ii++)
      aprime[ii]    = aprime[ii-1] + 0.5*dt*input[ii-1] +
          0.5*dt*input[ii-1];
      return aprime;
}


/* leaky integration */
float  * intfilt( float *input, int nt, float sampfrq, float rho)
{
    float *aprime;
    int ii;
    double dt = 0.5/sampfrq;

    aprime = ealloc1float(nt);
    memset((void *) aprime, 0 , nt*FSIZE);


    for (ii=1; ii<nt; ii++)
      aprime[ii]    = rho*aprime[ii-1] + dt*input[ii];
    return aprime;
};
```

Bibliography

[1] N. Bleistein, J.K. Cohen, and J. Stockwell. *Mathematics of multidimensional seismic Migration, Imaging and Inversion.* Springer–Verlag, 2000.

[2] N. K. Bose. *Digital Filters.* Elsevier Science Publishing Co. Inc, 1985.

[3] R. N. Bracewell. *The Fourier Transform and Its Applications.* McGraw–Hill, Inc., 1986.

[4] J. F. Claerbout. *Imaging the Earth's Interior*, pages 209–213. Blackwell Scientific Publications, 1985.

[5] A. V. Oppenheim, R. W. Schafer, and J.R. Buck. *Discrete-Time Signal Processing.* Prentice Hall, second edition, 1999.

[6] J. Tompkins, W. *Biomedical Digital Signal Processing: C–Language Examples and Laboratory Experiments for the IBM ®.* Prentice Hall, 2000.

[7] E Wegert and G. Semmler. Phase plots and complex functions a journey in illustration. *Notices Amer. Math. Soc.*, 58(6):768,780, 2011.

Appendices