Int. J. Appl. Math. Sci. Vol. **1**(2004), pp. 111-116

Explanation of Feynman's paradox concerning low-pass filters

A.G. Ramm
Mathematics Department
Kansas State University
Manhattan, KS 66506
ramm@math.ksu.edu

L. Weaver
Physics Department
Kansas State University
Manhattan, KS 66506
lweaver@phys.ksu.edu

Abstract In Feynman's lectures there is a remark about the limiting value of the impedance of an n-section ladder consisting of purely reactive elements (capacitances and inductances). The remark is that the limiting impedance $z = \lim_{n\to\infty} z_n$ has a positive real part. He notes that this is surprising since the real part of each z_n is zero, therefore it is impossible for the limit to have a positive real part. We explain the paradox, give the rate of convergence of the infinite sequence z_n, calculate the speed at which energy is propagated out into the infinite ladder, and comment on some of the literature on this topic.

1 Introduction

In Feynman's lectures [1] there is a remark about the limiting value of the impedance of an n-section ladder consisting of purely reactive elements (capacitances and inductances). The remark is that this limiting impedance $z = \lim_{n\to\infty} z_n$ has a positive real part. He notes that this is a paradox since the real part of each z_n is zero, therefore it is impossible for the limit to have a positive real part. A recent article [2] offered an explanation of this paradox based on the idea that real impedances have finite real parts and that we should define the limit as the iterated limit $\lim_{r\to 0}\lim_{n\to\infty}$. They attempted to show that this limit exists using the contraction mapping principle. Unfortunately their arguments do not give a correct and complete explanation of the paradox. Such an explanation is given apparently for the first time in our paper.

After introducing the problem in this section we give in section 2 a rigorous proof of convergence using only elementary arguments, and calculate the speed at which energy is propagated out into the infinite ladder. In section 3 we discuss the contraction mapping principle, explain some flaws in [2], and make some remarks.

Consider an n-section ladder (see Fig. 1).

If a section is added the impedance obeys the recurrence relation

$$z_{n+1} = Z_1 + 1/(1/Z_2 + 1/z_n) \tag{1}$$

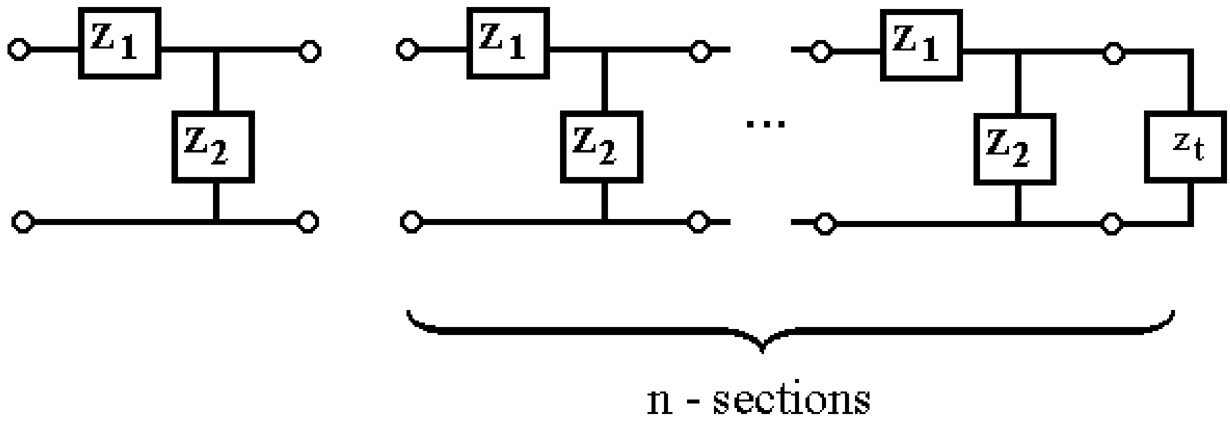

Figure 1: The network we discuss. In a real filter n is finite and the terminating impedance z_t is approximately z_+ (see text for notation).

and

$$z_1 = z_t, \tag{2}$$

where Z_1 and Z_2 are the impedances in each section, and z_t is the terminating impedance. In [2] $z_1 = Z_1 + Z_2$, and our theory covers this case.

The problem is to decide if the limit $\lim_{n\to\infty} z_n$ *exists, how it depends on* z_t*, and what is the rate of convergence to the limit.*

The example used by Feynman has $Z_1 = i\omega L, Z_2 = 1/(i\omega C)$. This is a low-pass filter: frequencies $\omega < \omega_c := 2/\sqrt{LC}$ are passed and frequencies above ω_c are rejected as $n \to \infty$. Feynman does not mention z_t explicitly but since he speaks of an infinite network containing only reactances, presumably $z_1 = Z_1 + Z_2$. He *implicitly assumes* that the limit $z = \lim_{n\to\infty} z_n$ exists. If this assumption holds, then z solves the equation

$$z = Z_1 + (Z_2^{-1} + z^{-1})^{-1} \tag{3}$$

and one easily finds its solutions:

$$z_\pm = Z_1/2 \pm \sqrt{Z_1^2/4 + Z_1 Z_2}. \tag{4}$$

He chooses the plus sign, a choice we discuss below. For $\omega > \omega_c = 2/\sqrt{LC}$ the square root is purely imaginary while for $\omega < \omega_c$ the square root is real. Now every term in the sequence z_n has zero real part, and for $\omega > \omega_c$ the real part of z is also zero, so in this case there is no paradox, but if $\omega < \omega_c$ the real part of z is positive and we have a *contradiction*, not a paradox. This shows that *the assumption that the limit exists for* $\omega < \omega_c$ *is wrong*: for $\omega < \omega_c$ *the limit does not exist.*

However Feynman was not making a blunder. His intuition, if correctly interpreted, led him to a correct conclusion. Namely, we prove that if $Z_1 = i\omega L + r$, $Z_2 = 1/(i\omega C) + r'$ with r, r' arbitrarily small positive real numbers, then the limit

$$\lim_{r,r'\to 0} \lim_{n\to\infty} z_n = z_+ \tag{5}$$

does exist for any $z_t \neq z_-$, and the rate of convergence is that of a geometric series. The choice r and r' non-negative is required for passive elements like L and C: they cannot create energy.

2 Existence of the Limiting Impedance

The notation is simplified by defining

$$p_n = z_n/Z_1, \quad p_\pm = z_\pm/Z_1, \quad \text{and} \quad t = Z_2/Z_1, \tag{6}$$

where $z_\pm$ are the solutions of equation (3). Then Eq. (1) becomes $p_{n+1} = 1 + tp_n/(t + p_n)$, the limiting value $p = p_+$ or p_- obeys $p^2 = p + t$, and $p_\pm = (1/2)(1 \pm \sqrt{1+4t})$. Define the branch of the square root by $\sqrt{1+4t} = a + ib$ with $a > 0$. This is always possible when $t \notin (-\infty, -1/4)$. Our first goal is to show that $\lim_{n\to\infty} p_n = p_+$. We will see that p_- is an unstable fixed point.

Theorem: *For $p_1 \neq p_-$ and $Re\sqrt{1+4t} > 0$ the sequence defined by*

$$p_{n+1} = 1 + tp_n/(t + p_n) \tag{7}$$

converges to p_+ at the rate of a geometric series.

Proof: First, note that if the sequence starts with $p_1 = p_-$ then every term $p_n = p_-$, so in this singular case the sequence does converge, but to the "wrong" impedance. So consider now $p_1 \neq p_-$.

It is easily checked that

$$\frac{p_{n+1} - p_+}{p_{n+1} - p_-} = \left(\frac{p_n t}{p_n + t} - \frac{p_+ t}{p_+ + t}\right) / \left(\frac{p_n t}{p_n + t} - \frac{p_- t}{p_- + t}\right) = \frac{p_n - p_+}{p_n - p_-} \cdot \frac{p_- + t}{p_+ + t} = \frac{p_n - p_+}{p_n - p_-} \cdot \frac{p_-^2}{p_+^2}. \tag{8}$$

Let $c_n := \frac{p_n - p_+}{p_n - p_-}$ and $\gamma := \frac{p_-}{p_+}$. Then Eq. (8) implies

$$|c_{n+1}| = |c_n||\gamma|^2, \tag{9}$$

and $|\gamma|^2 = [(a-1)^2 + b^2]/[(a+1)^2 + b^2]$ is less than 1 *if and only if* $a > 0$. Thus $|c_n| = |\gamma|^{2n}|c_1| \to 0$ as $n \to \infty$ and so $p_n \to p_+$. This shows that the limit $\lim_{n\to\infty} p_n$ *does exist* for all $p_1 \neq p_-$, it has positive real part, and the convergence is that of a geometric series. □

Remark 1: *If $p_1 = p_-$, then $p_n = p_-$ for all n, so the limit $\lim_{n\to\infty} p_n = p_-$ does exist in this case also. However, an arbitrarily small perturbation of the initial data $p_1 = p_-$ results in the limit $\lim_{n\to\infty} p_n = p_+$. This means that p_+ is the stable fixed point for the iterative process (7), while p_- is an unstable fixed point for this process.*

Remark 2: *There is a geometrical interpretation of the above proof which explains the reasons for formula (8) to hold. Namely, the mapping $p \to c(p) := (p - p_+)/(p - p_-)$ maps the complex plane onto itself, taking p_+ to the origin and p_- to the point at infinity. In the complex c-plane the map (8) is a dilation: $c_{n+1} = \gamma^2 c_n$. Thus the origin is a global attractor for $|\gamma| < 1$.*

Consider the low-pass filter discussed by Feynman. In this case $Z_1 = i\omega L + r$, $Z_2 = 1/(i\omega C) + r'$ with r, r' small and positive, so

$$\begin{aligned} 1 + 4t &= +A - i\epsilon, A > 0, \quad \text{for} \quad \omega > \omega_c \\ &= -A - i\epsilon, A > 0, \quad \text{for} \quad \omega < \omega_c. \end{aligned} \tag{10}$$

In these formulas $\epsilon = O(r + r')$ as $r, r' \to 0$, so ϵ is a small positive number, and $A > 0$ is $A = |1 - \omega_c^2/\omega^2|$. Choosing the branch of the square root with positive real part gives

$$\begin{aligned}\sqrt{1+4t} &= +\sqrt{A}-(i\epsilon/2)\sqrt{A}+O(\epsilon^2) \quad \text{for} \quad \omega>\omega_c\\ &= -i\sqrt{A}+(\epsilon/2)\sqrt{A}+O(\epsilon^2) \quad \text{for} \quad \omega<\omega_c.\end{aligned} \tag{11}$$

Here and below $\sqrt{A}>0$. Rewrite this in terms of the limiting impedance $z_+ = Z_1p_+ = (Z_1/2)(1+\sqrt{1+4t})$. For $\omega<\omega_c$ we find

$$z_+ = [(i\omega L+r)/2][1-i\sqrt{A}+(\epsilon/2)\sqrt{A}+O(\epsilon^2)] = (\omega L/2)\sqrt{A}+i\omega L/2+O(\epsilon). \tag{12}$$

The real part of z_+ is positive, consistent with the principle that the real part of the impedance of a passive network cannot be negative. We can now take the limit $\epsilon\to 0$ and get $\lim_{\epsilon\to 0} z_+ = (\omega L/2)\,(\sqrt{A}+i)$, which has a positive real part as was to be shown.

For $\omega>\omega_c$ we find

$$z_+ = [(i\omega L+r)/2][1+\sqrt{A}-(i\epsilon/2)\sqrt{A}+O(\epsilon^2)] = (i\omega L/2)(1+\sqrt{A})+O(\epsilon). \tag{13}$$

Again, the real part is positive (so it is physically reasonable), and we can take the limit $r, r'\to 0$ and find $\lim_{\epsilon\to 0} z_+ = (i\omega L/2)(\sqrt{A}+1)$. In this case, $\omega>\omega_c$, the impedance is purely imaginary. Taking the limit $\epsilon\to 0$ is equivalent to taking $r, r'\to 0$.

The requirement needed to make the mathematics rigorous is exactly the requirement that makes physical sense: the real part of a physically realistic passive impedance must be positive.

Once that is recognized one can describe the physical idealization of perfect impedances by taking the limit as these real parts go to zero.

We have treated a low-pass filter here, but a similar analysis applies to a high-pass filter in which the inductance and capacitance are interchanged.

Now that we know that the limits make sense we can discuss the infinite network further. If the infinite ladder is connected to a source, and the source is modulated in time, then the modulation propagates out into the ladder like a wave with well-defined group velocity. Thus energy can be delivered by the source out into the infinite network and the energy does not come back, it is absorbed: the infinite network acts like a "black box" whose impedance has a positive real part.

To see this, first use Kirchhoff's laws to show that $\tilde{V}_n(\omega) = (-\gamma)^n\tilde{V}_s(\omega)$ as done in [1], where $\gamma = p_-/p_+$ was used in the proof of the theorem. Then put $-\gamma = e^{i\delta(\omega)}$. Now describe the source voltage $V_s(t)$ by the modulated signal $V_s(t) = f(t)e^{i\omega_0 t}$, where the modulating function f is slowly varying. The Fourier transform of the source voltage is $\tilde{V}_s(\omega) = \tilde{f}(\nu)$, where $\nu := \omega-\omega_0$, and the spectrum of f vanishes outside a small neighborhood $|\nu|<\nu_0$ of the zero frequency. Then the time-dependence of the voltage on the n−th section of the ladder is $V_n(t) = e^{i\omega_0 t}\frac{1}{2\pi}\int_{|\nu|<\nu_0} d\nu e^{i\nu t+in\delta(\omega_0+\nu)}\tilde{f}(\nu)$. Compare this with the standard representation of a wave packet $g(x,t) := \frac{1}{2\pi}\int_{|\nu|<\nu_0} d\nu e^{i\nu t-ixk(\nu)}\tilde{f}(\nu)$, with group velocity $v := \frac{d\nu}{dk}|_{\nu=0}$. In our case the role of the variable x is played by parameter n and $k = -\delta(\omega_0+\nu)$, so $v = -1/\tau$, where $\tau := \frac{d\delta(\omega)}{d\omega}|_{\omega=\omega_0} > 0$. The quantity τ has the physical meaning of the time needed for the wave to propagate through one section of the ladder.

3 The Contraction Mapping Principle

In [2] the authors propose to apply the contraction mapping principle to prove the convergence of z_n for $\omega<\omega_c$, where

$$z_{n+1} = f(z_n), \quad z_1 = Z_1 + Z_2, \tag{14}$$

and

$$f(z) = Z_1 + 1/(1/Z_2 + 1/z). \tag{15}$$

They argue that the sequence converges to a fixed point $\zeta = f(\zeta)$ only if the mapping in equation (15) is a contraction mapping ([2], line 1 below (5)). This is wrong. Let us review the contraction mapping principle.

The contraction mapping principle says:

If there is a closed set D such that (a) $f(D) \subseteq D$ and (b) $|f(z') - f(z)| \leq q|z' - z|$ for all $z, z' \in D$ with q a constant, $0 < q < 1$, then there is in D a unique fixed point ζ of the map f, the iterative process $z_{n+1} = f(z_n)$, $z_1 = z_1 \in D$ converges for any initial approximation $z_1 \in D$: $\lim_{n\to\infty} z_n = \zeta$, and $\zeta = f(\zeta)$.

In [2] the authors do not specify a D that is mapped into itself by f nor do they check that their initial approximation z_1 generates a sequence that ever reaches D. They point out the latter gap in a footnote. They have simply calculated $|f'(\zeta)|$ for the fixed points ζ and claimed that only if $|f'(\zeta)| < 1$ will the sequence converge. This is false in general, (but, as we have proved, it is true for the linear fractional function with two distinct fixed points). A simple counterexample is $f(z) = z^2 + 1/4$. The sequence $z_{n+1} = f(z_n)$ converges to the fixed point $\zeta = 1/2$ for z_1 in $D = [0, 1/2]$ but $f'(1/2) = 1$, and although $f(D) \subseteq D$, f is not a contraction mapping because there is no $q < 1$ that satisfies condition (b) above for this set D. Further, the example $f(z) = \tan(z)$ shows that fixed points are possible with $|f'(\zeta)| \geq 1$, because at all the fixed points $\zeta_j = \tan \zeta_j$ we have $|f'(\zeta_j)| = |\sec^2 \zeta_j| \geq 1$.

In Section 2 we have proved convergence of the sequence in equation (1) without appeal to the contraction mapping principle. Our proof shows that if the conditions specified in section 2 are satisfied, then there exists a closed set $D_\epsilon = \{z : |z - z_+| \leq \epsilon\}$ such that for sufficiently small $\epsilon > 0$ the map f does map D_ϵ into itself and is a contraction. The initial approximation $z_1 = Z_1 + Z_2$ does not belong to this set, but our proof does show that for any $z_1 \neq z_-$ the approximations do eventually reach D_ϵ. This result the authors of [2] did not prove.

In conclusion we rephrase the remarks in section 2 in more precise language.

Remark 3: *There is a geometrical interpretation of the above proof which explains the reasons for formula (7) to hold. Namely, the mapping $p \to c(p) := S(p) = (p - p_+)/(p - p_-)$ maps the complex plane onto itself, taking p_+ to the origin and p_- to the point at infinity. In the complex c-plane the origin is a global attractor. The maps S, S^{-1}, and $f = [(1/t + 1)p + 1]/(p/t + 1)$ are linear fractional transformations. Such transformations form a group [3] that can be represented by 2×2 matrices (with determinant $+1$ in our case). Their composition is equivalent to multiplication of the corresponding matrices. These particular transformations are represented by the matrices*

$$S = \frac{1}{\sqrt{(p_+ - p_-)}} \begin{pmatrix} 1 & -p_+ \\ 1 & -p_- \end{pmatrix}, \; S^{-1} = \frac{1}{\sqrt{(p_+ - p_-)}} \begin{pmatrix} -p_- & p_+ \\ -1 & 1 \end{pmatrix}, \; f = \begin{pmatrix} 1/t + 1 & 1 \\ 1/t & 1 \end{pmatrix}. \tag{16}$$

Thus the transformation $p_{n+1} = f(p_n)$ is equivalent to $c_{n+1} = SfS^{-1}c_n$. The matrix D representing SfS^{-1} is easily calculated and is diagonal:

$$D = \begin{pmatrix} -p_-/p_+ & 0 \\ 0 & -p_+/p_- \end{pmatrix}. \tag{17}$$

The linear fractional transformation corresponding to the iteration $p_{n+1} = f(p_n)$ is then $c_{n+1} = \gamma^2 c_n$, $\gamma = p_-/p_+$, which converges for any initial approximation c_1, if $|\gamma| < 1$.

Remark 4: *Note that if $f(p) = 1+p/(p/t+1)$ then $f'(p) = t^2/(p+t)^2$, and $|f'(p_+)| = 1$ for $t = t_0 := -(\omega^2 LC)^{-1} < 0$. However, if $t = t_0 + i\eta$ with $\eta > 0$, then $|f'(p_+)| < 1$. Here $p_+ = [1+\sqrt{1+4t}]/2$. If f is a linear fractional map (in our case $f := [(1/t+1)p+1]/(p/t+1)$, and compared with the general case $f := (ap+b)/(cp+d)$ one has $a = 1+1/t, b = 1, c = 1/t, d = 1$, the determinant $ad-bc = 1$), then the condition $|f'(p_+)| < 1$, where p_+ is one of the two distinct fixed points of f, implies convergence of the iterations $p_{n+1} = f(p_n)$ for any initial approximation $p_1 \neq p_-$, as we have proved in Section 2. In other words, the point p_+ is the global attractor for the iterative process $p_{n+1} = f(p_n)$, $p_1 \neq p_-$. For the normalized linear fractional maps (that is, for the maps with $ad-bc = 1$)the trace $trace(f)$ of the matrix of the map, $trace(f) = 2 + t^{-1}$ in our case, defines the properties of the iteration process. If the trace $trace(f)$ is a real number and $|trace(f)| > 2$, then the map f is hyperbolic, if $|trace(f)| = 2$ then f is parabolic, and if $|trace(f)| < 2$ then f is elliptic. Parabolic maps are conjugate to translations, elliptic maps are conjugate to rotations, and hyperbolic maps are conjugate to positive dilations. If the trace $trace(f)$ of a normalized linear fractional map is a complex number (as in the case $r, r' > 0$ in Section 2), then the map f is called loxodromic, and it is conjugate to a complex dilation. A map, conjugate to f, is defined as the map $S^{-1}fS$, (see [3]) for more details).*

References

[1] R. P. Feynman, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1964) Vol. II, Chapter **22**, 12-15.

[2] H. Krivine and A. Lesne, Phase transistion-like behavior in a low-pass filter, Am. J. Phys. **71**, 31-33 (2003).

[3] Shapiro, J. H.,*Composition operators and classical function theory*, Springer-Verlag, New York, 1993.