MUS420/EE367A Lecture 12
Wave Digital Filters

Stefan Bilbao and Julius O. Smith III (jos@ccrma.stanford.edu)
Center for Computer Research in Music and Acoustics (CCRMA)
Department of Music, Stanford University
Stanford, California 94305

January 8, 2007

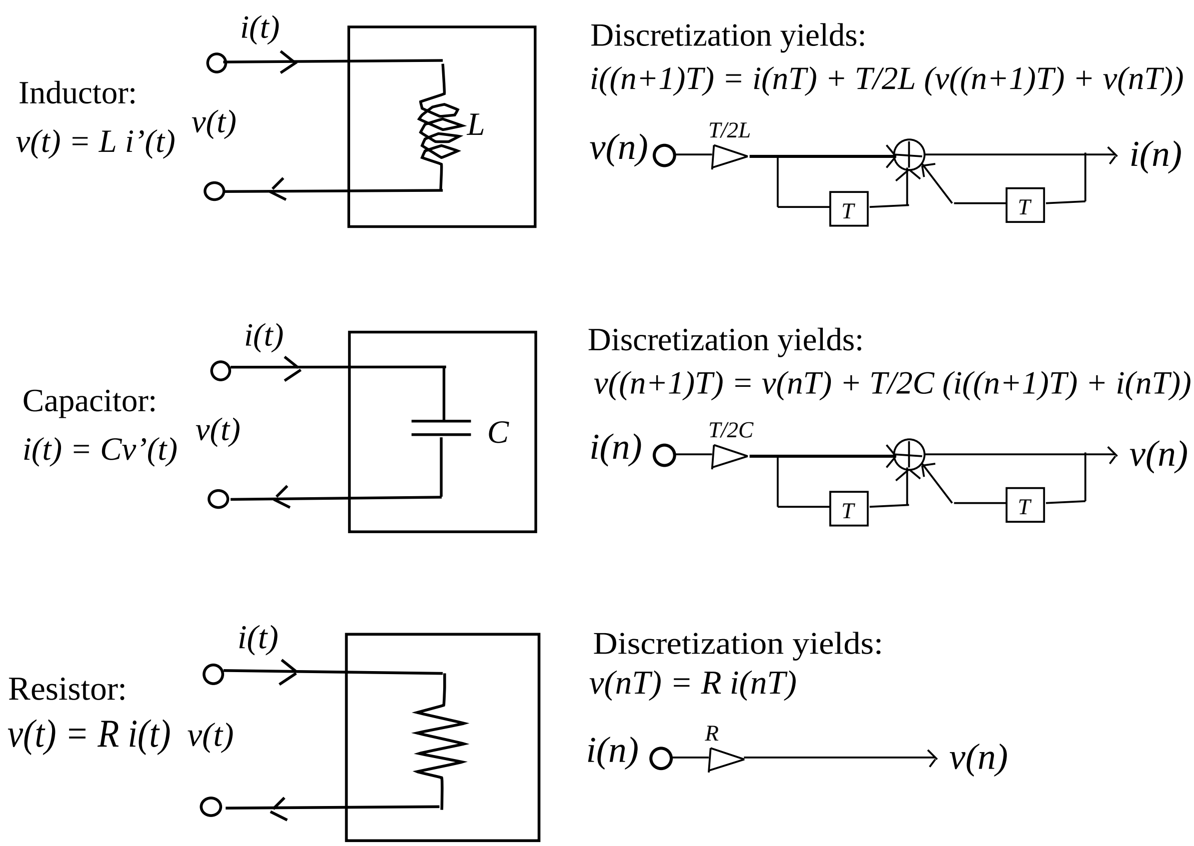
Outline:

- Finite Difference Schemes
- Delay Free Loops
- Wave Variables
- Wave Digital Inductor
- Bilinear Transform
- Scattering Junctions
- Physical Derivation of Wave Digital Filters
- Wave Digital Resonator Exercise
- Multidimensional Wave Digital Filters for solving PDEs

Wave Digital Filters

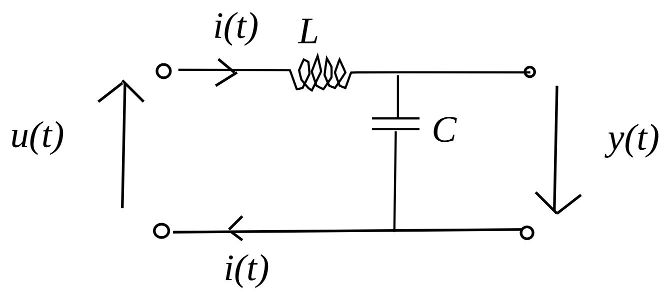
Wave digital filters were originally conceived as a way of discretizing an analog filter containing resistors, capacitors, inductors, transformers etc.

An alternative view would be that we are numerically integrating the set of ODEs which describes the analog circuit. Consider again the case of an inductor, which we will again discretize by the trapezoid rule for numerical integration (equivalent to the bilinear transform s-plane to z-plane mapping):



Delay-Free Loops

We might think, at this point, that we are done, because we can simply "replace" each RLC type circuit element by the derived signal flow path. Consider what happens with the discretization of the following simple LC filter:



Here, $u(t)$ is an input voltage, $y(t)$ the output voltage, $i(t)$ the current, and in addition, we will define $v(t)$ to be the voltage across the inductor. We thus have the three differential equations:

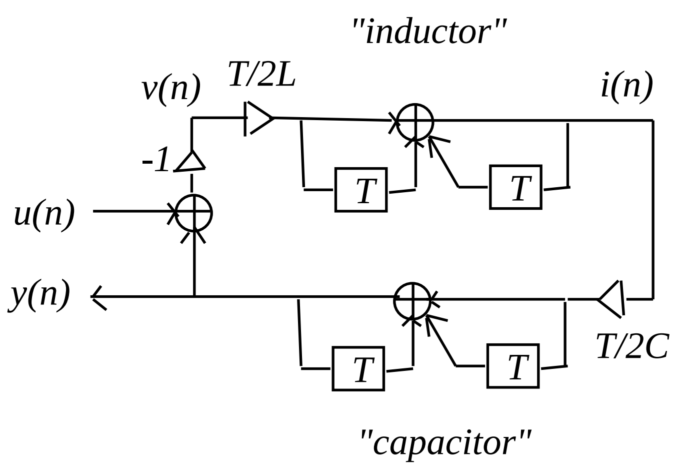
$$
\begin{aligned}
v &= L\frac{di}{dt} \\
i &= C\frac{dy}{dt} \\
v &= -u - y
\end{aligned}
$$

The first two come from the definitions of the inductor and capacitor respectively, the third from Kirchoff's voltage law.

Discretizing these equations according to the trapezoid rule yields:

$$\begin{aligned} i_n &= i_{n-1} + \frac{T}{2L}(v_n + v_{n-1}) \\ y_n &= y_{n-1} + \frac{T}{2C}(i_n + i_{n-1}) \\ v_n &= -(u_n + y_n) \end{aligned}$$

and this yields the following signal flow diagram:



Note however, the delay-free loop which prohibits this implementation from being realizable.

In other terms, connecting two *explicit* finite difference models has resulted in an *implicit* finite difference model.

Wave Variables

- In order to circumvent the problem of delay-free loops, Fettweis introduced *wave variables*:

$$\begin{aligned} a_n &= v_n + i_n R_0 \\ b_n &= v_n - i_n R_0 \end{aligned}$$

  where R_0 is an arbitrary parameter called a *port resistance* (for reasons we'll discuss later).

- In matrix notation, we readily see that this is a *linear transformation* of the state variables $\{v_n, i_n\}$:

$$\begin{bmatrix} a_n \\ b_n \end{bmatrix} = \begin{bmatrix} 1 & R_0 \\ 1 & -R_0 \end{bmatrix} \begin{bmatrix} v_n \\ i_n \end{bmatrix}$$

- Since the determinant of the two-by-two matrix is $-2R_0$, the transformation is non-singular provided $R_0 \neq 0$. We will additionally stipulate $R_0 > 0$.

- The inverse transformation is

$$\begin{aligned} v_n &= \frac{a_n + b_n}{2} \\ i_n &= \frac{a_n - b_n}{2R_0} \end{aligned}$$

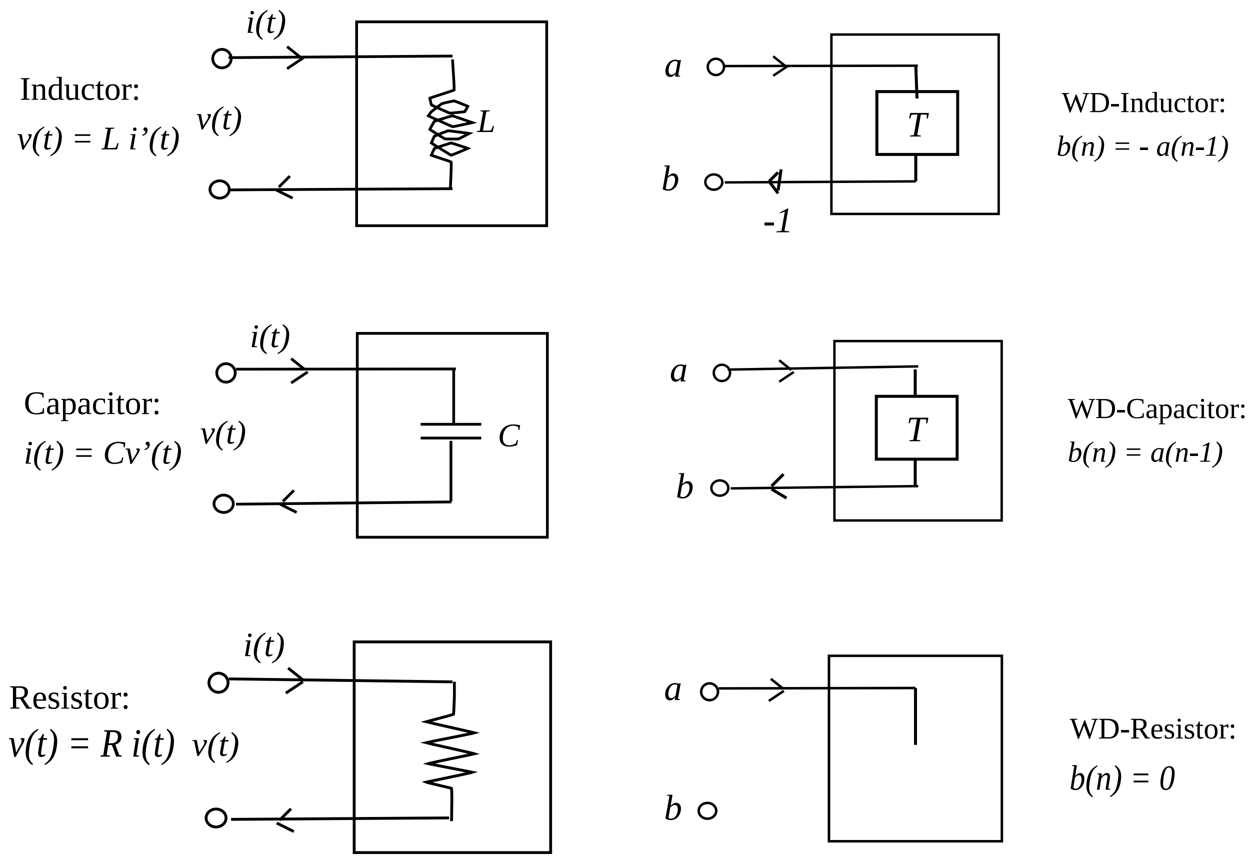
Derivation of the Wave Digital Inductor

It is instructive to see what happens when this change of variables in applied to the *inductor*.

1. We have the following *difference equation* for the inductor (using the trapezoidal rule of numerical integration, or bilinear transform, as you prefer):
$$i_n = i_{n-1} + \frac{T}{2L}(v_n + v_{n-1})$$
2. Perform the wave-variable substitution
$$v_n = \frac{a_n + b_n}{2} \qquad i_n = \frac{a_n - b_n}{2R_0}$$
to get
$$\frac{a_n - b_n}{2R_0} = \frac{a_{n-1} - b_{n-1}}{2R_0} + \frac{T}{2L}\left(\frac{a_n + b_n}{2} + \frac{a_{n-1} + b_{n-1}}{2}\right)$$
3. Now choose $R_0 = 2L/T$ to obtain
$$\begin{aligned} a_n - b_n &= a_{n-1} - b_{n-1} + (a_n + b_n + a_{n-1} + b_{n-1}) \\ &= 2a_{n-1} + a_n + b_n \end{aligned}$$
which further simplifies to
$$\boxed{b_n = -a_{n-1}} \qquad \text{(Wave Digital Inductor)}$$

Now there is no direct path from input to output.

In terms of *wave variables*, with simplest choices of the port resistances, we obtain the following wave digital filter elements (the elementary *one-ports*):

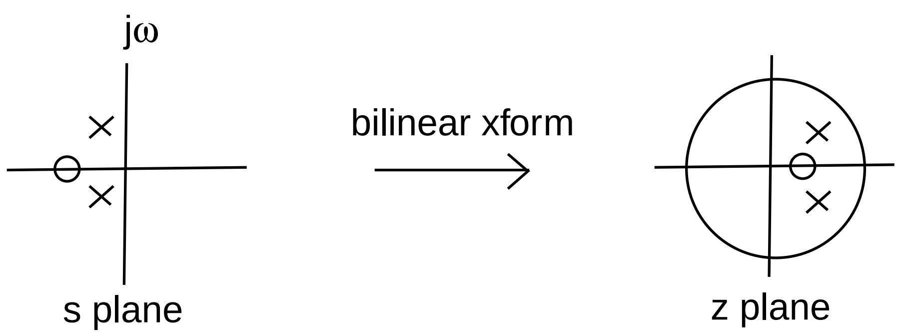


Elementary wave-digital one-ports. The port impedances for the wave-digital inductor, capacitor, and resistor (on the right) are defined as $2L/T$, $T/(2C)$, and R, respectively.

Note on the Bilinear Transform

Recall the mapping $s \longrightarrow \frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}$, which maps real frequencies to real frequencies. In fact, the mapping takes the RHP to the outside of the unit circle and the LHP to the inside. Thus:

- Under this particular bilinear transform, a stable minimum phase continuous time system is mapped to a stable minimum phase discrete time system.
- A function PR in the RHP is mapped to a function PR outside the unit circle.



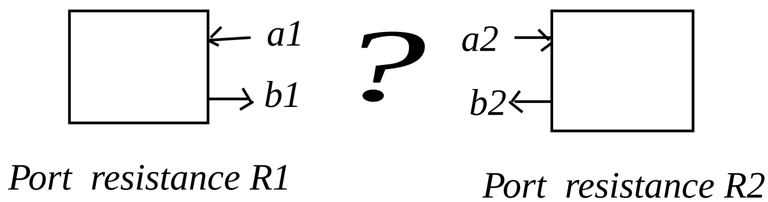
- Thus, in some sense this mapping *preserves passivity* or energetic properties of the original system.
- This is the fundamental reason why it is used by Fettweis et al. as the basis for filter design and numerical integration.

Building WDFs

- These wave digital one-ports may now be connected (via scattering junctions, to match impedances) in order to simulate analog circuits, or, equivalently, mechanical systems of masses, springs and dashpots.
- Since, in wave coordinates, there is no direct through path in any of the one-ports, delay-free loops cannot occur.
- Since “passivity” is preserved under the bilinear transform, the filter/numerical integrator is guaranteed stable.

Scattering Junctions

Suppose now that we would like to draw a new signal flow graph, using wave quantities. The problem now is that in general, the wave variables corresponding to a particular WD 1-port are scaled by $different$ port resistances. So how do we connect them?



We need to derive $scattering\ equations$...indeed, whenever we have an impedance change (even an artficial one such as this), we expect relection and transmission.

A Physical Derivation of Wave Digital Filters

- To each element, such as a capacitor or inductor, attach a length of transmission line at impedance R_0, and make it infinitesimally long. (Take the limit as the length of the transmission line goes to zero.)
  - The infinitesimal transmission line is *terminated* by the element.
  - The line impedance is *arbitrary* because it has been physically introduced.
  - If two such line-augmented elements are connected together by their transmission lines, scattering will clearly be induced at the junction in the usual way.
- Calculate the *reflectance* of the terminated line. That is, find the Laplace transform of the return wave divided by the Laplace transform of the input wave going into the line:
  - For a capacitor C (impedence $R_C(s) = 1/(Cs)$), we get the reflectance $S_C(s) = (R_C(s) - R_0)/(R_C(s) + R_0)$, which simplifies to

$$S_C(s) = \frac{1 - R_0 C s}{1 + R_0 C s}$$

– For an inductor L, we get $(Ls - R_0)/(Ls + R_0)$, or

$$S_L(s) = \frac{s - R_0/L}{s + R_0/L}$$

– For a resistor R, we get $(R - R_0)/(R + R_0)$, or

$$S_R(s) = \frac{1 - R_0/R}{1 + R_0/R}$$

– Note that both the capacitor and inductor reflectances are *stable allpass filters*, as they must be. Also, the resistor reflectance is always less than 1, no matter what line impedance R_0 we choose.

- Observe that there is a natural choice for each transmission-line impedance which will give us a normalized, universal reflectance for each element:

  – For the capacitor, $R_0 = 1/C \Rightarrow$

$$S_C(s) = \frac{1 - s}{1 + s}$$

  – For the inductor, $R_0 = L \Rightarrow$

$$S_L(s) = -\frac{1 - s}{1 + s}$$

  – And for the resistor, $R_0 = R \Rightarrow$

$$S_R(s) = 0$$

- Going to discrete time via the bilinear transform means making the substitution
$$s = c\frac{z-1}{z+1}$$
where c is some arbitrary positive constant, usually taken to be $c = 2/T$.
- Solving for z gives
$$z = \frac{1+s/c}{1-s/c}$$
- In this case, we see that setting $c = 1$ further simplifies our universal reflectances in the digital domain:
  - For the "wave digital capacitor" (or spring)
$$S_C\left(\frac{z-1}{z+1}\right) = z^{-1}$$
  - For the "wave digital inductor" (or mass)
$$S_L\left(\frac{z-1}{z+1}\right) = -z^{-1}$$
  - And for the "wave digital resistor" (or dashpot)
$$S_R\left(\frac{z-1}{z+1}\right) = 0$$
  as before in the continuous-time case.

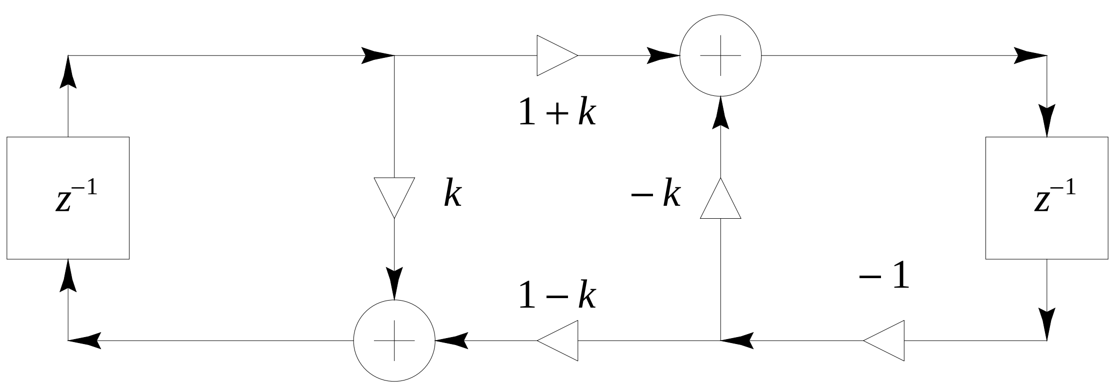
Equivalently, we may obtain the same results by setting $c = 2/T$ in the bilinear transform (which defines a frequency-scaling) and take the transmission-line (port) impedances to be instead $R_L = Lc = 2L/T$ for the inductor, and $R_C = T/(2C)$ for the capacitor (thereby compensating the frequency scaling).

A Wave Digital Resonator Exercise

Connect a wave digital capacitor and inductor together to form a second-order digital resonator consisting of

- a scattering junction in the middle,
- a unit-delay on the left (the capacitor), and
- a unit-delay and -1 gain on the right (the inductor).

You should be able to get a digital structure that looks like this:



Exercise, Cont.

For the exercise,

a) Find the reflection coefficient k of the induced scattering junction in terms of L and C.

b) Find the poles in terms of k.

c) Find the resonance frequency in terms of the sampling interval T and the reflection coefficient k.

d) Recall that an analog LC loop resonates at $1/\sqrt{LC}$, and relate these two resonance frequency formulas via the analog-digital frequency map $\omega_a = \tan(\omega_d T/2)$.

e) Show that the trig identity you discovered in this way is true.

This exercise verifies that the elementary "tank circuit" always resonates at exactly the frequency it should, according to the bilinear transform mapping.

Features of Wave Digital Filters (and Bilinear Transforms)

- We obtain an exact, one-to-one mapping of the frequency response $S(j\omega)$ (a reflectance) to the unit circle in the z plane, even though each first-order element is reduced to a mere unit delay with a possible sign flip (or to nothing but 0 in the case of a resistor/dashpot).

- One can show that a unit-sample delay in the digital version corresponds to an *exactly correct* phase-shift at the mapped analog frequency. The frequency mapping is, for this bilinear transform,

$$\omega_a = \tan\left(\frac{\omega_d T}{2}\right),$$

  where $\omega_a =$ analog radian frequency and $\omega_d =$ digital radian frequency.

- If it is possible to *warp* the frequencies of an input signal in the same way, that is, if we can perform the bilinear transform also on our analog input signal

$$x(t) \leftrightarrow X(s)$$

to produce a digital version

$$X\left(\frac{z-1}{z+1}\right) \leftrightarrow x_d(n),$$

then we can obtain *exact LTI processing of an infinite-bandwidth signal* inside a wave digital filter, no matter what sampling rate we choose!

- Unfortunately, this is usually not possible.
- **Exercise:** List the practical problems associated with trying to do something like this in practice, however approximate.

- Because the frequency axis is *warped*, we cannot expect nonlinear or time-varying systems to behave in discrete time as they did in continuous time:
  - Modulation sidebands land at "wrong" frequencies.
  - Harmonic relationships are destroyed $\Rightarrow$ "harmonic distortion" becomes inharmonic.

A quick look at solving PDEs via Multi-dimensional WDFs

Basic idea:

Replace a system of PDEs by a set of Kirchoff Loop Equations, each of which contains *multidimensional* (distributed) circuit elements.

Benefits:

- Associates a passive set of equations with a passive electrical circuit $\Rightarrow$ good stability properties.
- Algorithm will be explicit, parallelizable
- Extensions possible to non constant-coefficient, nonlinear problems, still retaining stability properties.

Drawbacks:

- Only works for PDEs derived from conservation relations (a small, but important class)
- Convergence may be slow.

Reference: Stefan Bilbao's thesis[1]

[1] `http://ccrma.stanford.edu/~bilbao/`

Distributed Nonlinearities in Music

- Shock waves in trombones
- Gongs
- Crashed cymbals
- Turbulence in Flutes
- Turbulence in the vocal tract

Distributed Modeling Example: 1-D wave equation

Let's look at the continuity equations for the 1-D acoustic tube, in the linearized, adiabatic case:

- Conservation of Mass:
$$\frac{\partial \rho}{\partial t} = -\frac{\partial (\rho u)}{\partial x}$$
- Conservation of Momentum:
$$\frac{\partial (\rho u)}{\partial t} = -c^2 \frac{\partial \rho}{\partial x}$$

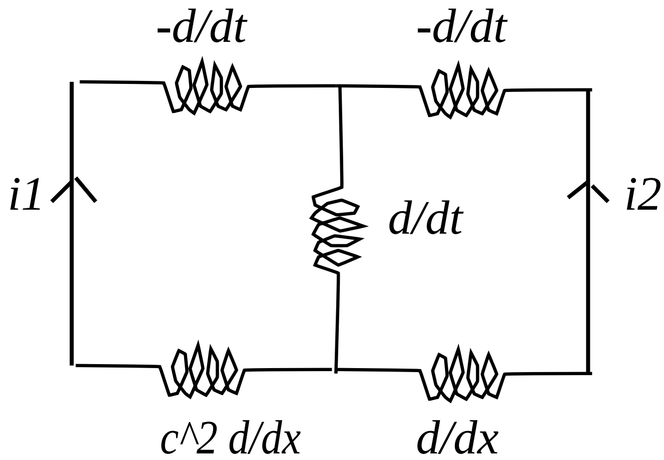
Now, let us rename density ρ by a variable i_1, and momentum density ρu by a variable i_2, which we will associate with currents:

$$\frac{\partial i_1}{\partial t} = -\frac{\partial i_2}{\partial x}, \frac{\partial i_2}{\partial t} = -c^2 \frac{\partial i_1}{\partial x}$$

and finally:

$$\frac{\partial i_1}{\partial t} + \frac{\partial i_2}{\partial x} = 0, \frac{\partial i_2}{\partial t} + c^2 \frac{\partial i_1}{\partial x} = 0$$

Now each $\frac{\partial}{dt}$ or $\frac{\partial}{dx}$ applied to a "current" such as i_1 or i_2 might be thought of as some sort of voltage corresponding to a generalized inductor (i.e., temporal or spatial). This is merely a way of making the jump to the following circuit representation of the equations:



- After a coordinate change, and impedance transformations (to remove negative inductances), we are left with a circuit that can be *discretized* according to WDF principles outlined above.
- MD-WDF one-ports are simple generalizations of their 1-D equivalents, except that internal delays may include spatial steps as well.
- Scattering is still memoryless (KCL equations depend only on voltage and current, not on the independent variuables).