A MEASURE-CONVOLUTION APPROACH TO ELEMENTARY SIGNALS AND SYSTEMS

Jeffrey O. Coleman

Naval Research Laboratory
Washington, DC

University of Maryland
Baltimore County

jeffc@alum.mit.edu

ABSTRACT

Linear, time-invariant (LTI) systems operating on and producing complex measures instead of ordinary time functions provide a convenient unification and generalization of ordinary discrete-time and continuous-time LTI systems. Further, a development of this approach paralleling that of a typical elementary signals and systems course appears reasonably straightforward, building only on a prerequisite minimal working familiarity with Lebesgue integration and the Radon-Nikodym derivative.

1. INTRODUCTION

The continuous-time and discrete-time signals and systems of signal processing lead to a variety of convolution expressions to deal with their various combinations. Roberts and Mullis [1], for example, comment that

> "There are 6 combinations of 2 signal classes. For each one of these 6 combinations, one can define a convolution operation which involves one signal from each of the two signal classes."

Indeed, even with the three common aperiodic convolutions,

$$y(n) = x(n) \star h(n) = \sum_k x(k)h(n-k)$$
$$y(t) = x(n) \star h(t) = \sum_k x(k)h(t-kT)$$
$$y(t) = x(t) \star h(t) = \int x(\tau)h(t-\tau)\,d\tau$$

one often feels obliged to derive a result for a specific combination, for example, the associativity result

$$\begin{aligned} &((a(n) \star b(n)) \star p(t)) \star q(t) \\ = \; &a(n) \star (b(n) \star (p(t) \star q(t))) \end{aligned}$$

But all the while there is the feeling that the derivation ought not to be necessary, that it that ought to simply follow from some more-general result. This paper represents a first attempt at moving towards a unified representation of signals, systems, and convolutions with the aim of combining the cases and eliminating such redundant derivations. The approach involves going back to the fundamentals of LTI system representation and using complex measures.

The author is with NRL, but this work relates more to his adjunct faculty appointment in the UMBC Dept. of CS and Electrical Engineering.

Certain uses of measures in such contexts is certainly well established. Riesz representation theorems [2] suggest representing a linear, time-invariant (LTI) system having input $x(t)$ and output $y(t)$ by an integral of the form

$$y(t) = \int x(t-\tau)\,d\phi(\tau) = \int x(t-\tau)\,dU(\tau)$$

where the complex measure ϕ in the Lebesgue integral or the associated unit-step response $U(t)$ in the Stieltjes integral captures the system behavior. If ϕ is absolutely continuous with respect to Lebesgue measure λ and $h(t)$ is the complex Radon-Nikodym derivative $d\phi/d\lambda$, then the familiar convolution $y(t) = \int x(t-\tau)h(\tau)\,d\tau$ results. The notion of signal representation by a measure is also well estabished. For example, the same integrals with x representing the system and ϕ or U representing the input signal might typically be used in a mathematical development of LTI systems driven by white noise [3].

This paper goes one step further and develops the simultaneous (deterministic) representation of both signals and systems as measures in order to unify discrete and continuous time in a clean way, especially at the boundary between the two, when one of either signal or system is continoustime and the other is discrete-time. The approach is to sketch the development of these ideas informally and nonrigorously but in a way that could be made credible and appealing for, say, first- or second-year graduate students in engineering. The idea is to require no more background in analysis than can be comfortably presented in the early, background portion of a graduate course. (Only a little of that background is outlined here.)

As a generalization then, the present approach represents an alternative to the theory of distributions [4] and/or generalized functions, the usual route to unification of discrete and continuous time and to the formal development of the Dirac delta function. That usual approach is the stronger for dealing with differentiation and differential equations, but the measure-theoretic approach examined here is better able to handle fractal-type signal structures. An example based on

Conf. on Information Sciences and Systems
Baltimore, MD, March 1999

the Cantor distribution function is briefly sketched at the end of the paper.[1]

The next section is the heart of the paper. It develops the notions of measure input/output LTI systems, measure convolution and its elementary properties, and the key mathematical concepts and results needed to tie measure convolution to ordinary convolution. The third section actually makes that connection to those ordinary convolutions and is followed by a discussion of some things one might conclude from all this.

2. LTI SYSTEMS ON COMPLEX MEASURES

After defining a linear, time-invariant measure system this section outlines the development of its representation by a convolution of measures. The interesting part turns out to be the definition of time invariance, which must be handled with care to cover cleanly the cases that will eventually correspond to the usual continuous-time and discrete-time systems. Unless explicitly stated, all measures in this development (and paper) are complex measures on the Lebesgue measurable subsets of the reals and all constants and functions are complex valued. Complex measures, as in $\hat{x}$, will generally be distinguished from functions of the same name by a hat.

To begin, suppose operator $\mathcal{L}$ represents a system with measures for input and output. The system is *linear* if $\mathcal{L}(\alpha\hat{x}+\beta\hat{y}) = \alpha\mathcal{L}(\hat{x}) + \beta\mathcal{L}(\hat{y})$ always holds.

2.1. System Time Invariance with Respect to a Measure

If delayed measure $\hat{z}_\tau$ is defined by $\hat{z}_\tau(A) = \hat{z}(A-\tau)$ in terms of the translation by τ of arbitrary measurable set A, then of course the temptation is to call the system time invariant if $\mathcal{L}\hat{z}_\tau = (\mathcal{L}\hat{z})_\tau$ always holds. But in some cases "always" is too strong. The usual time-invariance notion for discrete-time systems, for example, involves only integer translations. In general the values of τ for which such translation invariance will turn out to be important are those values relevant to some enclosing integral over τ (recalling that an enclosing discrete-time sum is just an integral with respect to a measure on sets of integers). This next definition formalizes the notion. Keep in mind that $\mathcal{L}\hat{z}$ is a measure and $(\mathcal{L}\hat{z})(A)$ is that measure applied to measurable set A.

Definition. *A system $\mathcal{L}$ is time invariant with respect to measure $\hat{x}$ if for every measure $\hat{z}$,*

$$\int_B (\mathcal{L}\hat{z}_\tau)(A)\, d\hat{x}(\tau) = \int_B (\mathcal{L}\hat{z})_\tau(A)\, d\hat{x}(\tau)$$

for all measureable sets A and B.

Inside the integral, $\mathcal{L}\hat{z}_\tau$ has had exactly the same effect as $(\mathcal{L}\hat{z})_\tau$. This definition amounts to the statement that if $\mathcal{L}\hat{z} = \hat{y}$, then effectively, $\mathcal{L}\hat{z}_\tau = \hat{y}_\tau$ inside integrals $d\hat{x}(\tau)$. Since it is so for an arbitrary $\hat{z}$ and A, it can only be because $\hat{x}$ attributes no measure to the set of τ values for which translation invariance would fail. Further, the formal definition has an exact if informal differential equivalent[2]

$$d(\mathcal{L}\hat{z}_\tau)(s)\, d\hat{x}(\tau) = d(\mathcal{L}\hat{z})_\tau(s)\, d\hat{x}(\tau)$$

Recall that a complex measure $\hat{x}$ is *absolutely continuous* with respect to positive measure ν, denoted $\hat{x} \ll \nu$, if the support of $\hat{x}$ is contained in the support of ν, that is, if $\nu(A) = 0$ always implies $\hat{x}(A) = 0$. Suppose, for example, that σ is counting measure on the integers. Then any measure $\hat{x} \ll \sigma$ can be thought of as a discrete-time signal. Using absolute continuity and this next simple result, time invariance does not have to be handled separately for each measure.

Proposition. *If $\hat{x} \ll \nu$ and $\mathcal{L}$ is time invariant with respect to ν, then $\mathcal{L}$ is also time invariant with respect to $\hat{x}$.*

Roughly, if ν has no measure on the set of τ where translation invariance fails, then the same is true of $\hat{x}$. A more formal approach requires the Radon-Nikodym derivative.

Recall that the *Radon-Nikodym derivative* of measure $\hat{x}$ with respect to positive measure ν, denoted $d\hat{x}/d\nu$, is a concept that applies only when $\hat{x} \ll \nu$ (along with other conditions on the measures that we will not discuss here). Then we say that the (ν-integrable) function $h = d\hat{x}/d\nu$ if

$$\hat{x}(A) = \int_A h(t)\, d\nu(t)$$

for every measurable set A. This integral relationship is fully equivalent to the differential relationship $d\hat{x} = h(t)\, d\nu(t)$. A probability density function of a scalar random variable, for example, is the Radon-Nikodym derivative of a probability measure (the one induced on the reals by the random variable) with respect to Lebesgue measure on the real line.

To address the proposition then, suppose that $x = d\hat{x}/d\nu$ and write

$$\begin{aligned} d(\mathcal{L}\hat{z}_\tau)(s)\, d\hat{x}(\tau) &= d(\mathcal{L}\hat{z}_\tau)(s) x(\tau)\, d\nu(\tau) \\ &= d(\mathcal{L}\hat{z})_\tau(s) x(\tau)\, d\nu(\tau) \\ &= d(\mathcal{L}\hat{z})_\tau(s)\, d\hat{x}(\tau) \end{aligned}$$

Since time invariance with respect to counting measure σ actually amounts to time invariance with respect to every discrete-time measure, it might even be fair to consider σ to represent in some sense discrete time itself and, likewise, to consider Lebesgue measure λ to represent continuous time. We will see that this notion holds up well throughout the development.

[1] NRL's Abraham Shultz offered the observation that the fractal structure of many wavelet problems makes them application candidates as well.

[2] It is so because each integral in the definition is a complex measure on measurable rectangle $A \times B$, which rectangles form a π-system generating the Borel subsets of the plane [5].

2.2. System Characterization by Measure Convolution

Let the term *impulse* be identified with the Dirac measure δ that represents a unit point mass at the origin. Call complex measure $\mathcal{L}\delta$ the *impulse response* of system $\mathcal{L}$ and here denote it by $\hat{h}$. For any measure $\hat{x}$ on any measurable set A,

$$\hat{x}(A) = \int \delta(A-\tau)\, d\hat{x}(\tau) = \int \delta_\tau(A)\, d\hat{x}(\tau)$$

because the integrand is equal to the indicator (characteristic) function $1_A(\tau)$ of A. If the system displays an appropriate sort of continuity (analogous to the usual BIBO property in elementary signals and systems) then taking $\hat{x}$ as the system input will result in a system output $\mathcal{L}\hat{x}$ characterized by a similar integral with δ_τ replaced by the system's response to it, so

$$(\mathcal{L}\hat{x})(A) = \int (\mathcal{L}\delta_\tau)(A)\, d\hat{x}(\tau)$$

If $\mathcal{L}$ is time invariant with respect to $\hat{x}$, then inside the integral $\mathcal{L}\delta_\tau = \hat{h}_\tau$, and

$$(\mathcal{L}\hat{x})(A) = \int h(A-\tau)\, d\hat{x}(\tau)$$

This relationship is a *convolution of measures* and is economically represented by the notation $\mathcal{L}\hat{x} = \hat{h} \star \hat{x}$. This was derived assuming that the system was time invariant with respect to $\hat{x}$, but it is easy to show a converse result, that a system $\mathcal{L}$ defined by this relationship is necessarily time invariant with respect to $\hat{x}$.

2.3. Elementary Measure-Convolution Properties

It is suprising how many elementary proofs involving convolution of measures use Fubini's theorem to reverse the order of integration where one of the integrals has arisen from the relationship $\hat{x}(A) = \int 1_A(\tau)\, d\hat{x}(\tau)$. The proof of the commutativity of convolution is a good example. In the spirit of an elementary presentation, many steps are included:

$$\begin{aligned}
(\hat{x} \star \hat{y})(A) &= \int \hat{x}(A-t)\, d\hat{y}(t) \\
&= \int\int 1_{A-t}(s)\, d\hat{x}(s)\, d\hat{y}(t) \\
&= \int\int 1_{A-s}(t)\, d\hat{y}(t)\, d\hat{x}(s) \\
&= \int \hat{y}(A-s)\, d\hat{x}(s) \\
&= (\hat{y} \star \hat{x})(A)
\end{aligned}$$

That commutativity property can now be used in a detailed associativity proof. Fubini's theorem is again invoked midway through.

$$\begin{aligned}
(\hat{x} \star (\hat{y} \star \hat{z}))(A) &= ((\hat{y} \star \hat{z}) \star \hat{x})(A) \\
&= \int (\hat{y} \star \hat{z})(A-t)\, d\hat{x}(t) \\
&= \int\int \hat{y}(A-t-s)\, d\hat{z}(s)\, d\hat{x}(t) \\
&= \int\int \hat{y}(A-s-t)\, d\hat{x}(t)\, d\hat{z}(s) \\
&= \int (\hat{y} \star \hat{x})(A-s)\, d\hat{z}(s) \\
&= ((\hat{y} \star \hat{x}) \star \hat{z})(A) \\
&= ((\hat{x} \star \hat{y}) \star \hat{z})(A)
\end{aligned}$$

The third basic property is the simple observation that convolution with an impulse is an identity operation:

$$(\delta \star \hat{x})(A) = (\hat{x} \star \delta)(A) = \int \hat{x}(A-t)\, d\delta(t) = \hat{x}(A)$$

Finally, many familiar Fourier-transform properties can be proved from this definition of the Fourier Transform $\hat{X}$ of a measure $\hat{x}$:

$$\hat{X}(A) = \int \left(\int_A e^{-i2\pi f t} df \right) d\hat{x}(t)$$

2.4. Shift Invariance of One Measure with Respect to Another

To show in Section 3 that the Radon-Nikodym derivatives of input, output, and impulse-response measures correspond to ordinary inputs, outputs, and impulse responses in discrete- or continuous-time, we will need the fact that sometimes a measure is invariant under certain input translations. For example, Lebesgue measure λ on the real line is invariant under all translations, so that $\lambda = \lambda_\tau$ for all real τ. But consider counting measure σ on the integers. Clearly, $\sigma = \sigma_\tau$ does not hold for all real τ, though it does hold for integer τ. In general the values of τ for which we need such translation invariance are those values relevant to an enclosing integral over τ. A time-invariance notion for measures strikingly parallel to the time-invariance notion already developed for systems formalizes this notion.

Definition. *If μ and $\hat{y}$ are measures, we say that μ is shift invariant with respect to $\hat{y}$ if*

$$\int_B \mu_\tau(A)\, d\hat{y}(\tau) = \mu(A)\hat{y}(B)$$

for all measureable sets A and B.

The effect of $\mu_\tau(A)$ inside the integral is the same as $\mu(A)$, because $\hat{y}$ attributes no measure to the set of τ values for which translation invariance would fail. Counting measure on the integers is, for example, shift invariant with respect to counting measure on the even integers, though the converse is false. This formal definition of shift invariance is fully equivalent to the more usable form

$$d\mu_\tau(s)\, d\hat{y}(\tau) = d\mu(s)\, d\hat{y}(\tau)$$

Proposition. *If $\hat{y} \ll \nu$ and μ is shift invariant with respect to ν, then μ is also shift invariant with respect to $\hat{y}$.*

This is because

$$\begin{aligned} d\mu_\tau(s)\, d\hat{y}(\tau) &= d\mu_\tau(s)\, y(\tau) d\nu(\tau) \\ &= d\mu(s)\, y(\tau) d\nu(\tau) \\ &= d\mu(s)\, d\hat{y}(\tau) \end{aligned}$$

3. CONTINOUS-TIME AND DISCRETE-TIME SIGNALS AND SYSTEMS

The relation of measure convolution to ordinary convolution is through the Radon-Nikodym derivative. Here a derivative property for measure convolution provides the basis of this relationship, following which some specific relationships to continuous-time and discrete-time are examined.

3.1. The Radon-Nikodym Derivative of a Convolution

Consider measure convolution $\hat{x} \star \hat{y}$ and suppose one of the measures $\hat{x}$ and $\hat{y}$ represents an input signal while the other represents a system impulse response. The Radon-Nikodym derivative of system output $d(\hat{x} \star \hat{y})/d\mu$, where μ typically represents the discrete-time or continuous-time "domain," can then be obtained easily if certain conditions are met. Suppose in particular that $\hat{x} \ll \mu$ with μ shift invariant with respect to $\hat{y}$. Let $x = d\hat{x}/d\mu$ for notational convenience. Then (using substitution $t = s - \tau$ to obtain the third line)

$$\begin{aligned} (\hat{x} \star \hat{y})(A) &= \int \hat{x}(A - \tau)\, d\hat{y}(\tau) \\ &= \int \int 1_{A-\tau}(t) x(t)\, d\mu(t)\, d\hat{y}(\tau) \\ &= \int \int 1_A(s) x(s - \tau)\, \underbrace{d\mu_\tau(s)\, d\hat{y}(\tau)}_{d\mu(s)\, d\hat{y}(\tau)} \\ &= \int_A \int x(s - \tau)\, d\hat{y}(\tau)\, d\mu(s) \\ &= \int_A (x \star \hat{y})(s)\, d\mu(s) \end{aligned}$$

expressing the result finally in terms of a convolution of a function with a measure. This reveals a basic result:

Proposition. *When $\hat{x} \ll \mu$ with μ shift invariant with respect to $\hat{y}$,*

$$\frac{d}{d\mu}(\hat{x} \star \hat{y}) = \frac{d\hat{x}}{d\mu} \star \hat{y} \qquad (1)$$

paralleling the familar signal-processing notion that under many circumstances, differentiation commutes with other LTI operations.

3.2. Relating Measure Convolution and Ordinary Convolution

Here continous time and discrete time are represented by particular positive measures with respect to which Radon-Nikodym derivatives may be taken. Let λ denote Lebesgue measure on the reals, as before, and let positive measure Σ^T denote counting measure on $T\mathbb{Z}$. Let $\hat{z} = \hat{x} \star \hat{y}$, and note that in each of these four cases

$\hat{x} \ll$	$\hat{y} \ll$
λ	λ
Σ^T	Σ^T
Σ^T	Σ^{MT}
λ	Σ^T

the positive measure on the left is shift invariant with respect to the measure on the right and hence with respect to $\hat{y}$ as well. We will see that each row then corresponds to a particular specialization of (1) to an ordinary convolution relationship in terms of x and z, the Radon-Nikodym derivatives of $\hat{x}$ and $\hat{z}$ with respect to the domain-representation measure on the left, and y, the Radon-Nikodym derivative of $\hat{y}$ with respect to the domain-representation measure on the right. As a common first step, (1) becomes $z = x \star \hat{y}$, but the meaning of this then needs examination on a case-by-case basis.

Case 1: Two continuous-time signals. With functions x, y, and z representing the Radon-Nikodym derivatives of $\hat{x}$, $\hat{y}$, and $\hat{z}$ with respect to Lebesgue measure λ, the relationship is

$$z(t) = \int x(t - \tau)\, y(\tau)\, d\tau$$

an ordinary continuous-time convolution.

Case 2: Two discrete-time signals. If $x(n)$, $y(n)$, and $z(n)$ represent the Radon-Nikodym derivatives, evaluated at time nT, of $\hat{x}$, $\hat{y}$, and $\hat{z}$ with respect to counting measure Σ^T, the relationship becomes

$$z(n) = \sum_k x(n - k)\, y(k)$$

an ordinary discrete-time convolution.

Case 3: Two sample rates. With $z(n)$ and $x(n)$ representing the Radon-Nikodym derivatives, evaluated at time nT, of $\hat{z}$ and $\hat{x}$ with respect to counting measure Σ^T, and with $y(n)$ denoting the Radon-Nikodym derivative, evaluated at time nMT, of $\hat{y}$ with respect to counting measure Σ^{MT}, the relationship becomes

$$z(n) = \sum_k x(n - kM)\, y(k)$$

which in digital signal processing represents the discrete-time convolution of x with y after the latter has been zero-interpolated by M.

Case 4: One discrete-time and one continuous-time signal. With $z(t)$ and $x(t)$ representing the Radon-Nikodym derivatives, evaluated at time t, of $\hat{z}$ and $\hat{x}$ with respect to Lebesgue measure λ, and with $y(n)$ denoting the Radon-Nikodym derivative, evaluated at time nT, of $\hat{y}$

with respect to counting measure Σ^T, the relationship becomes

$$z(t) = \sum_k x(t - kT)\, y(k)$$

a mixed convolution that in communication theory is typically described as as pulse amplitude modulation (PAM) of sequence $y(k)$ with pulse shape $x(t)$.

4. DISCUSSION AND CONCLUSIONS

From the four cases at the end of the last section, we see that essentially all aperiodic convolutions of ordinary functions in continuous and discrete time are fully equivalent to convolution of the measures associated to those functions through integration. We can see in fact that if we identify every continuous-time function $x(t)$ with a measure defined by

$$\hat{x}(A) = \int_A x(t)\, dt$$

and every discrete-time function $x(n)$ at sampling interval T with a measure defined by

$$\hat{x}(A) = \sum_{\{k:\, kT \in A\}} x(k)$$

then ordinary LTI systems (even involving zero-interpolation) and PAM modulators are characterized indirectly by measure convolution. (And in fact a signal or impulse response may be mixed and have components of both types.) Some of the mathematical forms even have a familiar look. A finite-impulse-response (FIR) filter, for example, has an impulse-response measure $\hat{h} = \sum_k h_k \delta_{kT}$ defined by

$$\hat{h}(A) = \sum_k h_k \delta(A - kT)$$

where here δ represents a unit point measure at the origin.

At first glance measure convolution may seem like simply an elaborate way to formalize the notion of the Dirac delta function, the use of which formally permits discrete-time signals or components of signals to be represented in continuous time. But it can in fact handle certain signals and systems that other methods (even the theory of distributions) cannot.

Consider, for example, an LTI system in continuous time whose unit-step response $U(t)$ is the Cantor distribution function from probability theory [5]. This function is continuous but has a zero derivative almost everywhere, so it is not possible to obtain the system output $y(t)$ as an ordinary convolution of the system input $x(t)$ with the derivative of the unit-step function. At the very least, the Stieltjes integral

$$y(t) = \int x(t - \tau)\, dU(\tau)$$

is required (equivalent to the convolution of an input function and an impulse-response measure). But it gets much worse when two such systems are cascaded and we desire to characterize the composite system by a single convolution. Ordinarily, by the associativity of convolution, we convolve two impulse-response functions to arrive at the impulse response of the cascade. What do we do here? Convolution of the impulse-response measures is fully defined and handles the situation as naturally as can be hoped for. But without a tool of this type we are lost.

5. REFERENCES

[1] Richard A. Roberts and Clifford T. Mullis, *Digital Signal Processing*, Addison-Wesley, Reading, MA, 1987.

[2] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, third edition, 1987.

[3] Eugene Wong and Bruce Hajek, *Stochastic Processes in Engineering Systems*, Springer-Verlag, New York, 1985.

[4] W. Rudin, *Functional Analysis*, McGraw-Hill, New York, second edition, 1991.

[5] P. Billingsley, *Probability and Measure*, Wiley, New York, second edition, 1986.