DYNAMICAL THEORY FORMALISM FOR ROBUST MODELING OF DAMPED, UNDAMPED, AND NONLINEAR OSCILLATORY SIGNALS

Irina Gorodnitsky, Senior Member, IEEE

University of California, San Diego
La Jolla, CA, USA 92093-0515
igorodni@cogsci.ucsd.edu

ABSTRACT

The paper explores a novel framework for signal representation based on dynamic information in a signal that is well suited for robust analysis of low SNR signals and extraction of time-varying features. The method is derived from dynamical theory but formulated in a basic parameter estimation paradigm. Modeling the *changes* in data provides a compact depiction of time-variant and invariant information plus features related to data dynamics. The method also provides strong noise mitigation properties even when noise statistics is poorly understood. The signal processing formulation supplies a connection between the time-delay and the Fourier domains. This connection helps us bridge non-linear dynamical and signal processing theories and brings a powerful novel tool to signal analysis at large. The experiment is presented using a speech sample from the TIMIT database.

Index Terms— Signal processing, Time-varying systems, Volterra series, Feature extraction, Speech processing

1. INTRODUCTION

Many excellent methods exist for analysis of stationary data contaminated by Gaussian noise. The situation is less satisfactory when signals are time-variant, both transient and long-scale nonstationary. Equally challenging can be analysis of signals degraded by high and 'non-standard' (e.g. non-Gaussian and non-stationary) noise.

This paper introduces a domain for signal analysis where we exploit the signal *derivative*. This leads to a method that appears to be well suited for the scenarios described above. The method also offers very simple means for non-linear data analysis. The suggested domain is a low-order non-uniform time-delay embedding [1] to which we map the *derivative* of the data vector. While the embedding concept originated in the context of nonlinear dynamics, this theory is exploited here in a signal processing context. Namely, the framework is established which allows us to compute time-delay domain features from noisy data via a basic parameter estimation paradigm. Interest in non-linear methods has been limited in the signal processing (SP) community because classical nonlinear techniques tend to be computationally complex and are notoriously extremely sensitive to even minute random data perturbations. In contrast, the proposed method is highly robust and can be used without any reference to nonlinear dynamics theory. One of the contributions of the paper is establishing a connection between the Fourier (FD) and the proposed domains. Another contribution is establishing a link between nonlinear dynamics theory and the field of signal processing.

It is important to realize that embedding theory does not imply that data must be generated by a non-linear system. Rather it can be applied to data generated by any deterministic system, linear or nonlinear. The property that embedding theory provides which is exploited here is that we can represent highly dynamically complex but deterministic signals in a recurrence relationship. This recurrence relationship is guaranteed to exist, under fairly loose conditions, in a space that is spanned by a sufficiently large number of vectors composed of past observations of data.

Following pioneering contributions in [3], I define a signal model in the time-delay domain as a projection of the derivative of a stochastic observable into a finite order attractor, i.e. the manifold. The key attribute here is the parameter estimation framework through which the dimension and the embedding parameters, including the delays, are found directly from the data. Modeling the data derivative rather than the data vector itself gives rise to several non-trivial properties. We find a drastic reduction in the model orders that need to be used to represent even highly complex signals. All the dynamic models, including the simplest one containing a single delay term, describe an infinite set of outputs of damped and undamped harmonic oscillations plus a limited set of overdamped oscillations. The models with two delay terms can describe more complex signals, including those with nonlinear features. Thus even very basic dynamical models, with as few as two to four parameters, fit many of the data classes one typically expects to encounter in SP. The model universality precludes errors due to model misspecification for those classes of data. Strong noise mitigation properties which are the result of projection onto a low-order manifold further contribute to the method's robustness. Overall, we find the method to be highly robust in the case of very short data records and low SNR.

One interesting interpretation of the delay domain models is in the context of timing sequences or spike trains underlying biological cell signaling. In this interpretation, a single time-delay, whose value adapts to the changes in data over time is analogous to an interspike interval. Timing codes are widely believed to be responsible for the admirable efficiency with which biological systems can encode time-variant information. When we consider dynamical delay models, a single delay can encode an infinite set of oscillatory functions. Extensive investigations, both in-vivo and via modeling, have attempted to uncover the underlying principle of the biological code. However, the possibility that spike trains may encode the *changes in the signal* has not been explored. Such encoding would be highly efficient for transmitting time-varying information and almost universal in its modeling property in the sense that it can represent a large set of functions. The topic is not discussed further here, except to note that the proposed domain and its estimation algorithm have been termed correspondingly as the Interval Domain (ID) and Interval Domain Estimation Algorithm (IDEA).

2. INTERVAL DOMAIN DERIVATION

We consider a one-dimensional observation vector $\mathbf{x}(t)$ to be an output of a l-dimensional system. We also define a D-dimensional *time-delay reconstructed* state vector at time t (see [1]) as $\mathbf{x}_t = \{x(t), x(t \leq \tau_1), \ldots, x(t \leq \tau_{D-1})\}$, where $\tau_i,\ 1 \leq i \leq D \leq 1$ are delays. According to [1], under fairly generic conditions there exists a map between the l-dimensional system and a D-dimensional state vector $(D > 2l+1)$ that is a $C^k, k < l$ diffeomorphism, also called an embedding. It follows that we can express the *derivative* of the data vector as a smooth function of the state vector. I show in the next section that the choice to represent the derivative in the time-delay domain significantly expands the informational content that can be described by very simple models.

The form of the function into which the derivative would map in the delay coordinates is unknown but it can always be defined in the general form as a functional expansion with unknown parameters which are constrained by the data. Volterra kernel functions, which represent a hierarchy of interactions among the past and present states of the system, provide a good modeling choice. Specifically, we can write

$$\begin{aligned}\dot{\mathbf{x}}(t) = \quad & \mathbf{g}_0 + \sum_{\tau=0}^{\infty} \mathbf{g}_\tau \mathbf{x}_\tau + \sum_{\tau_1=0}^{\infty}\sum_{\tau_2=0}^{\infty} \mathbf{g}_{\tau_1,\tau_2}\mathbf{x}_{\tau_1}\mathbf{x}_{\tau_2} + \lll + \\ & \sum_{\tau_1=0}^{\infty}\sum_{\tau_2=0}^{\infty} \lll \sum_{\tau_q=0}^{\infty} \mathbf{g}_{\tau_1,\tau_2,\ldots,\tau_q}\mathbf{x}_{\tau_1}\mathbf{x}_{\tau_2}\ldots\mathbf{x}_{\tau_q}\ldots \quad ,\end{aligned} \tag{1}$$

where bold face letters denote vectors and $\mathbf{x}_{\tau_i} = \mathbf{x}_{t-\tau_i}$ is a state vector at time $t \leq \tau_i$. Note that we can generally omit the term $\mathbf{g}_0$ if we expect the signals to be bounded. The difference between Eq. (1) and a standard model is, of course, in the use of the derivative $\dot{\mathbf{x}}(t)$. Successful recovery of an embedding from generic data is governed by considerations covered in [2]. Particularly relevant to us is that the choice of the delay vectors is critical and that delays should be allowed to be non-uniform. It follows that rather than taking the classical approach of choosing delays and the model dimension ad hoc, it is preferable to estimate the embedding parameters directly from the derivative.

As already reviewed, a key advantage we gain over simply expanding the data vector is in drastic reduction in the model order that we can use. We find that even the simplest linear one-delay models, of which there are two,

$$\dot{\mathbf{x}} \quad = \quad a\mathbf{x}_\tau \tag{2}$$

$$\dot{\mathbf{x}} \quad = \quad a_0\mathbf{x} + a_1\mathbf{x}_\tau\ , \tag{3}$$

can be sufficient for many SP applications. These two models will be studied in detail here. We refer to the first model as the *basic* ID model and to the second as the general linear one-delay ID model.

The extension of this method to data contaminated by random noise is straightforward. In this case, exact estimate of a smooth invertible function from the derivative is not possible. Instead, we *force* the diffeomorphism through projection of the derivative into a low-order ID model by estimating the model parameters to within some error, the standard procedure in SP. The low-order models, such as those in Eqs. (2) and (3) can be estimated using a number of methods for solving underdetermined problems, which range from sparse optimization to global searches.

3. ANALYSIS OF THE ONE-DELAY ID MODELS

To understand properties of ID models, it is necessary to understand the data classes fitted by these models. Note that ID models are equivalent to autonomous Delay Differential Equations (DDE). Hence, the task here can be formulated as finding solution spaces of the corresponding DDE. Here we study signal classes described by the one-delay models in Eqs. (2) and (3). Through the process I show how the time and frequency domain signal information (amplitude and frequency) are represented in these models.

Solving Delay Differential Equations analytically is notoriously difficult except for some restricted cases. The essential difference in how individual solutions are defined in ODE and DDE is in the initial value problem. In ODE, the initial value problem is given by a point. In a DDE, the initial value problem is defined by a function. Hence, unlike an ODE, a DDE posses an infinite solution space, a fundamental property which explains how the simplest ID model can accommodate the large span of data classes.

Pure sinusoid: Before considering the infinite solution space of Eqs. (2) and (3), it will greatly benefit the discussion to examine how pure sinusoids are represented in the ID. Consider sinusoid $x(t) = A\sin(\omega t + \tau)$ with frequency $f = w/2\tau$. This function creates a one-dimensional limit cycle solution in the delay embedding space. Therefore, we expect the simplest model that fully describes its derivative to be a single delay model Eq. (2). Substituting the sinusoid signal into Eq. (2) we get $\frac{\omega}{a}\cos(\omega t + \tau) = sin(w(t \leq \tau) + \tau)$. There is only one possible choice for τ for which this equality can hold:

$$\tau_n = \frac{(2n \leq 1)\tau}{2\omega}, \qquad n \in I^+, \tag{4}$$

with the corresponding coefficient $a = \leq\omega$ for odd n and $a = \omega$ otherwise. As can be expected since the delay identifies recurrences in the data, τ_n is inversely related to the frequency. The multiple values of τ_n in Eq. (4) are analogous to the harmonics in the frequency domain. The smallest value $\tau = \frac{\tau}{2\omega} = \frac{1}{4f}$, corresponds to a quarter wavelength and is inversely proportional to the 4th harmonic. For consistency, we refer to the delay which corresponds to one full cycle as the *fundamental delay* $\tau_f = \frac{1}{f}$. In the case of a pure tone, the absolute value of the coefficient a is equal to the angular frequency. It is defined unambiguiously, unlike τ.

Note that since $\tau \in I^+$, only an approximation to the frequency that is the closest to the inverse integer τ value can be found in practice. Thus the accuracy of the method is bounded by the sampling rate and data upsampling is one way to obtain a finer τ resolution. This highlights an important point regarding utility of this method. IDEA is not a high precision algorithm that would make an automatic choice for analysis of pure stationary signals with low to moderate SNR. Its contributions are in improved robustness in high SNR and in efficient representation of time-varying signals which is shown next.

We now characterize the entire solution spaces of linear one-delay DDE shown in Eqs. (2) and (3). The derivations are presented first and results are summarized at the end.

As discussed, the solution space of a DDE is defined by the boundary value problem where the initial condition must be defined as a function $x(t) = u(t)$ on the $t \in [\leq\tau, 0]$ interval, called the *pre-interval*. Similar to ODEs the general nontrivial solutions to Eq. (3) have the form $x = B\mathrm{e}^{\tau t}$, which yields the following nonlinear, transcendental *characteristic equation*

$$\tau \leq a_0 \leq a_1\mathrm{e}^{-\tau\tau} = 0\,. \tag{5}$$

Characterizing roots τ_p of this transcendental equation can be nontrivial. Observing Eq. (5) for the different relationships between a_0,

a_1, and τ, we can see even before obtaining a formal solution that it admits an infinite number of complex roots. These roots give rise to an infinite set of oscillatory solutions.

The analytic solution for the case $a_0 = 0$ was given in [4]. Here we use a different approach to solve Eqs. (2) and (3) based on the concept of the Lambert function. The Lambert function, also known as the Omega function or product log [5], is a multivalued function $W(z)$ defined by the equation $W(z)\mathrm{e}^{W(z)} = z$, $z \in \mathcal{C}$. Multipling both sides of Eq. (5) by $\tau e^{a_0\tau}$, it is straightforward to show its roots for the two models can be written respectively as

$$\tau = \frac{1}{\tau}W(\leq a\tau) \tag{6}$$

$$\tau = \frac{1}{\tau}W(\leq a\tau\mathrm{e}^{a_0\tau}) \leq a_0\,. \tag{7}$$

In both cases, the roots of Eq. (5) are the values of the Lambert function scaled and shifted along the real axis (or simply scaled in the case of Eq. (6)).

The Lambert function is multivalued with infinitely many branches in the complex plane, conventionally labeled as $W(z)_k,\ k = 0, \leq 1$, etc. There are two cases where in addition to the infinitely many complex branches, there exist either two real branches $W_0 : \leq 1 \leq W_0 \leq 0$ and $W_{-1} : W_{-1} \leq \leq 1$ defined on the interval $[\leq 1/\mathrm{e} \leq z < 0)$ or one real branch $W_0 : W_0 > 0$ on the interval $[0, \inf)$. At point $z = \leq 1/\mathrm{e}$ both branches come together and, in effect, present a single real solution $W = \leq 1$.

The real versus complex $W(z)_k$ define fundamentally different forms of solutions. Since superposition of any of the individual solutions is also a solution, we can write the corresponding general solutions for (2), (3) as

$$x(t) = \sum_{k=-\infty}^{\infty} B_k\,\mathrm{e}^{\frac{1}{\tau}W(-a_1\tau)t} \tag{8}$$

$$x(t) = \sum_{k=-\infty}^{\infty} B_k\mathrm{e}^{\frac{1}{\tau}W(-a_1\tau\mathrm{e}^{a_0\tau})t}\mathrm{e}^{a_0 t}\,. \tag{9}$$

The scaling B_k is defined by the preshape function, which incorporates information on the amplitude of the signal. The two solution spaces in Eqs. (8) and (9) appear to be closely related. Both contain an infinite set of oscillatory intrinsic modes $\mathrm{e}^{\frac{1}{\tau}W_k(\cdot)t}$ defined by the complex $W_k(\cdot)$ and which describe the characteristic spectrum for the corresponding models. There are also three nonoscillatory modes related to the real roots. The difference between the two solution spaces is in the appearance of the second time-dependent exponential term in (9). The two exponential terms in Eq. (9), one primarily oscillatory and the other always nonoscillatory, describe distinctly different behaviors and have independent time scales. The real $e^{a_0 t}$ can be interpreted as modeling amplitude variations. If we consider the scenarios where estimation is done adaptively over short time windows, this term can be used to model amplitude modulations. Moreover, under the condition $|\dot{B}_k/B_k| << f_k$, where is f_k denotes frequency, amplitude modulation and phase modulation have similar presentation. Hence, the term could capture phase modulation in some types of signals. This means phase modulations can be identified even though constant phase information is lost in an embedding.

To understand the specific form of the oscillatory modes above we consider complex roots $\tau_k = r_k + jw_k$. The oscillatory solutions defined by $x = B\mathrm{e}^{\tau t}$ can be written simply as

$$x(t) = \sum_{k=-\infty}^{\infty} \mathrm{e}^{r_k t}\,(B_{1k}\sin(w_k t) + B_{2k}\cos(w_k t)). \tag{10}$$

Using substitutions $C_k = (B_{1k}^2 + B_{2k}^2)^{1/2}$, $sin(v_k) = B_{2k}/C_k$, $cos(v_k) = B_{1k}/C_k$ we rewrite the above as

$$x(t) = \sum_{k=-\infty}^{\infty} C_k\mathrm{e}^{r_k t}\,(\sin(w_k t + v_k)), \tag{11}$$

where v_k is an angular phase. We can see now that the data classes represented by one-delay ID models include all damped (and growing) and undamped harmonic oscillations. In addition, as explained above, model (3) also admits more complex waveforms, including modulated harmonic oscillations.

We next derive the relationships between the ID parameters and the time and frequency domain features of data. We consider an individual oscillatory component of the solution and drop the subscript k for convenience. Substituting $\tau = r + jw$ into the characteristic equation (5) we have $\tau = a_0 + a_1 e^{-r\tau}(\cos(\omega\tau) \leq i\sin(\omega\tau))$ and separating the real and imaginary parts after some manipulation:

$$\begin{aligned} a_0 &= r + \omega\cot(\omega\tau) \\ a_1 &= \leq\frac{\omega e^{r\tau}}{\sin(\omega\tau)}\,. \end{aligned} \tag{12}$$

This system can be solved numerically to yield a family of oscillatory solutions for each choice of $\{a_0, a_1, \tau\}$. A numerically more convenient formulation of the same is

$$\begin{aligned} &r = a_0 \leq \omega\cot(\omega\tau) \\ &\ln\left(\leq a_1\frac{\sin(\omega\tau)}{\omega}\right) \quad \leq a_0\tau + \omega\tau\cot(\omega\tau) = 0. \end{aligned} \tag{13}$$

The later gives a transcendental equation purely for w that can be solved numerically. Other types of expressions linking the features of the individual domains may be useful. For example, we can derive from Eq. (12)

$$\tau = \frac{1}{w}\tan^{-1}\frac{w}{a_0 \leq r}\,, \tag{14}$$

which gives a more explicit link between τ and the frequency and decay/growth amplitude information of a damped/forced sinusoidal oscillation $e^{rt}\cos(\omega t)$.

We can easily derive similar results for the basic ID model by setting $a_0 = 0$ in the above equations. The general form of the equations in not altered in this case, so that the relationships between the features of the two domains are similar although not identical. In particular, when $a_0 = 0$, Eq. (14) becomes $\tau = 1/w\tan^{-1} \leq w/r$. From this we can evaluate how a one-dimensional limit cycle in the embedding space is related to the two parameters of a harmonic oscillator. In particular, as the damping coefficient grows, $r >> w$, this expression is dominated by the first term of its Taylor series expansion $\tau = \leq 1/r$. As one can see, τ becomes primarily affected by the damping coefficient. In the limit, we have $w = 0$ and the roots are purely real. Solutions corresponding to the real roots can be viewed as outputs of an overdamped (overforced) oscillator. Since there can be up to two negative and one positive real τ, one-delay ID models can describe at most two overdampled oscillator outputs or one overforced oscillator output in any given data.

We can write the most general form of the solution then as

$$x(t) = B_0\,\mathrm{e}^{r_0 t} + B_1\,\mathrm{e}^{r_1 t} + \sum_{k=2}^{\infty}\mathrm{e}^{r_k t}\,(C_k\sin(w_k t + v_k)),\ C_k \neq 0 \tag{15}$$

where one or both coefficients $B_{0,1}$ assume non-zero values on the interval defined by $\leq a_1\tau\mathrm{e}^{a_0\tau} < 1/\mathrm{e}$. Note the change in the limits of the summation in Eq. 15 which does not affect the general result.

The interval where real $r_{0,1}$ exist can be partitioned further into sections that contain respectively one or two negative real $r_{0,1} < 0$, which define decaying exponentials, or one real $r_0 > 0$ and the other $r_1 \in \mathcal{C}$, which defines a single growing exponential. These subintervals can be easily found from the information given above on behavior of the Lambert function.

The key property of one-delay ID models uncovered by the analysis is their compact representation of any combination of transient and stationary harmonic oscillations. This property is summarized in the following theorem.

Theorem 1 *Any function $\mathbf{x}(t)$ composed of any number of harmonic oscillations (damped and undamped) may be expressed in the form of linear one-delay ID models $\dot{\mathbf{x}} = a\mathbf{x}_\tau$ and $\dot{\mathbf{x}} = a_0\mathbf{x} + a_1\mathbf{x}_\tau$.*

Analysis of higher order linear and nonlinear DDE shows that complexity of their solution spaces increases with an increase in the degrees of freedom. However, all DDE share a core subspace of solutions describing an infinite set of fundamental waveforms that are combinations of outputs of general undamped, damped and forced oscillators. The set of functions which can be fit by these combinations is quite large. Hence even the simplest ID models are quite universal in being able to describe many types of data.

Such universality is one reason behind the method's robustness. Classical parameter estimation methods essentially depend on the correct specification of the underlying model. Errors and approximations in models typically lead to inconsistent parameter estimates and compromised results. Since ID models are exact for many signals we expect to see, we circumvent the errors due to model misspecification in many cases. The absence of such errors is one reason for the observed robustness of DDE in modeling very short data. This ability, in turn, extends applicability of IDEA to adaptive processing over short frames of nonstationary data.

In comparison with FD, the Interval Domain extends considerably the classes of signals that can be efficiently represented. The information available to a user in the ID is in the form of compact features associated with amplitude and frequency of stationary and transient oscillations. The compact representation extends to data which are considerably complex or broadband. For such and other data, the features condensed into a sparse ID basis may reveal signal information which may not be apparent otherwise. Moreover, the novelty of the representation is not simply in its compactness. It can be shown (analysis omitted here) that in addition to regular moments of data, the ID extracts what may be called *dynamic moments*, which are cross-products of various powers of the signal and *its derivative*. How to exploit information carried in these moments is not clear at this point, but it is clear that they reveal dynamic aspects not accessible by standard SP methods.

In summary, the novel IDEA method provides three key properties: efficiency, robustness, and the ability to accommodate time-variant and even nonlinear signals within the efficient framework. Modeling the derivative in a delay domain was originally proposed by Kadtke [3], but the mathematical underpinnings, including the key capabilites of this method listed above, were not realized. These capabilities usher in many potential uses in compression, coding, recognition, classification, synthesis, and data mining applications.

4. EXPERIMENTS: PITCH ESTIMATION

Due to limited space we go directly to analysis of realistic data where we consider the problem of pitch estimation. Speech is an example of piecewise stationary or semi-stationary data containing abrupt transitions between segments which are often quite different, e.g. transitions between harmonic and nonharmonic phonemes. Due to this nonstationarity, speech must be analyzed over fairly short frames which leads to poor performance in high noise.

The example here uses a two-delay model $\dot{\mathbf{x}} = a_1\mathbf{x}_{\tau_1} + a_2\mathbf{x}_{\tau_2}$, which was not explored above due to limited space. From experience, this model provides slightly more robust estimates of pitch for certain phonemes than the one-delay linear ID models, although how significant the differences in estimates may be is still being investigated. The pitch is represented in the larger of the two delays of this model, which we will denote by τ_2.

The TIMIT database, sentence SX132, female speaker FAEM0 was chosen for the test. The parameters of the two-delay linear model were estimated for different levels of white noise added to the signal. In the interest of space, only τ_2 information is plotted. Fig. 1 shows the estimated τ_2 for the noise free case, converted to the frequency scale by the transform f_s/τ_2, where $f_s = 16000Hz$ is the sampling rate. Plotted on the same graph is pitch estimate F_0 computed using the ESPS method obtained from the Snack library for Linux platforms (`http://www.speech.kth.se/SNACK`). ESPS integrates normalized cross-correlation pitch candidate generation with dynamic programming to select the optimal pitch track. Both ESPS and IDEA were run with settings of 25 msec and 50.4 msec respectively for the frame spacing (framelength in SNACK) and the window size. The 107Hz lower bound was used with both methods and 400Hz upper bound was used for ESPS method. The ESPS method automatically calculates probability of voicing and returns pitch values for what is deemed to be the voiced speech segments and 0 otherwise. Similarly, I used 400Hz upper bound cut-off for IDEA output to separate voiced from unvoiced speech. Fig. 1 shows very close agreement between the two methods. IDEA pitch values are slightly higher, likely due to τ_2 being integer valued. The gap in performance becomes evident at negative SNR (results not plotted due to limited space). Specifically, at -5dB SNR the ESPS algorithm does not detect voiced speech in most of the sentence while IDEA is able to track pitch in parts of the sentence.

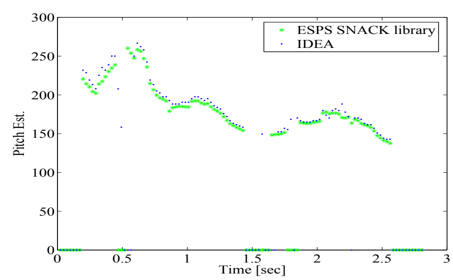

Fig. 1. Pitch estimated with IDEA and the ESPS (Snack library) algorithms for the TIMIT database sentence SX132, speaker FAEM0.

5. REFERENCES

[1] F. Takens, "Detecting strange attractors in turbulence," *Lecture Notes in Mathematics*, vol. 898, pp. 366, 1981.

[2] P. Grassberger, T. Schreiber, and C. Schaffrath, "Nonlinear time sequence analysis," *Int. J. Bifurc. Chaos*, vol. 1, pp. 521, 1991.

[3] J. Kadtke, "Classification of highly noisy signals using global dynamical models," *Phys. Lett. A*, vol. 203, pp. 196, 1995.

[4] C. E. Falbo, "Analytic and numerical solutions to the delay differential equation $y'(t) = \tau y(t \leq \tau)$," in *Joint meeting of the No and So Cal sections of the MAA*, 1995.

[5] E.M. Wright, "Solution of the equation $ze^z = a$," *Proc. Roy. Soc. Edinburgh*, vol. 65, pp. 193–203, 1959.