A Theoretical Explanation for the Efficiency of Generalized Harmonic Wavelets in Engineering and Seismic Spectral Analysis

Michael Beer[1], Afshin Gholamy[2], and Vladik Kreinovich[3]
[1]Institute for Risk and Reliability
Leibniz Universität Hannover
Hannover, Germany, beer@irz.uni-hannover.de
[2]Department of Geological Sciences
[3]Department of Computer Science
[2,3]University of Texas at El Paso
El Paso, TX 79968, USA
afshingholamy@gmail.com, vladik@utep.edu

Abstract

Wavelets of different shapes are known to be very efficient in many data processing problems. In many engineering applications, the most efficient shapes are shapes of a generalized harmonic wavelet, i.e., a wavelet of the shape $w(t) = t^a \exp(b \cdot t)$ for complex b. Similar functions are empirically the most successful in the seismic analysis – namely, in simulating the earthquake-related high-frequency ground motion. In this paper, we provide a theoretical explanation for the empirical success of these models.

1 Formulation of the Problem

Fourier transforms: a brief reminder. In signal processing, an important tool is Fourier transform, where a signal is represented as linear combination of sine and cosine waves $\sin(\omega \cdot t)$ and $\cos(\omega \cdot t)$ corresponding to different frequencies ω. The possibility for such representation can be traced to the famous Newton's prism experiment, when he showed that any light passing through a prism decomposes into a linear combination of monochromatic lights (from red to violet) – which are exactly sines and cosines.

The sinusoidal *infinite* periodic waves, whose frequencies do not change with time are indeed very efficient in analyzing *stationary* processes, i.e., processes whose statistical characteristics do not change with time. In particular, these waves are very useful in *spectral analysis*, i.e., in describing phenomena that include processes corresponding to different frequencies and periods. Examples of

such phenomena include sea waves, wind storms, vibrations in civil engineering structures, etc.

Wavelets: need to go beyond Fourier transforms. Many real-life processes are *non-stationary*. In particular, many engineering and seismic processes are non-stationary. During a short time interval, we can often safely ignore the changes in their statistical characteristics, but as we consider larger and larger time intervals, we are no longer able to ignore this change. Thus, locally, we can still describe the signal as a linear combination of periodic waves, but this representation can no longer be expanded to the whole real line.

So, to properly represent non-stationary signals, we need to use functions which are not periodic. These functions should instead be limited to a bounded time interval. Such time-bounded wave-like functions are known as *wavelets*.

Wavelets have indeed been successfully applied in many engineering applications.

Which wavelets are the most efficient in engineering applications: empirical analysis. There are many different types of wavelets. A comparative analysis of the efficiency of different types of wavelets – described in [3, 4, 7, 12, 13, 14] – shows that in many engineering applications, the most efficient wavelets are so-called *generalized harmonic wavelets* first introduced in [8, 9, 10], i.e., linear combinations of the functions $t^{-1} \cdot \exp(\mathrm{i} \cdot \omega \cdot t)$ corresponding to different values ω (here, as usual, $\mathrm{i} \stackrel{\text{def}}{=} \sqrt{-1}$).

Comment. In real-valued terms, harmonic wavelets are linear combinations of the functions $t^{-1} \cdot \sin(\omega \cdot t)$ and $t^{-1} \cdot \cos(\omega \cdot t)$. Both these functions tend to 0 as t goes to $+\infty$ or to $-\infty$. Thus, these functions are indeed localized.

From the mathematical viewpoint, we have a minor problem with the values corresponding to $t = 0$: since we cannot divide by 0, strictly speaking, the above functions are not defined when $t = 0$.

For the function $t^{-1} \cdot \sin(\omega \cdot t)$, this is not a serious problem, since this expression has a finite limit (equal to ω) when $t \to 0$. By using this limit as the value of the function when $t = 0$, we get a continuous bounded function – well-suited for describing continuous processes whose intensity is usually bounded by some constant (we are not talking about explosion-type processes when the values of some quantities can suddenly experience a drastic increase).

For the function $t^{-1} \cdot \cos(\omega \cdot t)$, however, the limit at $t \to 0$ is infinity. To make it bounded, we therefore cannot consider any such function by itself, we need to consider linear combinations of such functions, combinations arranged in such a way that infinities cancel each other and the limit becomes finite.

Similar functions are empirically most efficient in seismic analysis. Interestingly, similar functions, namely, functions of the type

$$w(t) = t^a \cdot \exp(b \cdot t),$$

turn out to be the most empirically successful in seismic analysis; see, e.g., [2, 5, 6, 11]. Specifically, to gauge the effect of a possible earthquake on an

engineering structure, it is necessary to be able to adequately simulate the corresponding ground motion. This simulated ground motion $g(t)$ is usually obtained by the following multi-stage procedure:

- first, we generate a white noise $n(t)$;
- then, we apply an appropriate shaping window $w(t)$ to the white noise, resulting in the function $x(t) \stackrel{\text{def}}{=} \int w(s) \cdot n(t-s)\, ds$;
- after that, we apply the Fourier transform to the function $x(t)$, resulting in $\widehat{x}(\omega)$;
- we multiply the resulting Fourier transform $\widehat{x}(\omega)$ by an appropriate function $s(\omega)$, to make sure that the spectrum of the resulting signal $\widehat{m}(\omega) \stackrel{\text{def}}{=} s(\omega) \cdot \widehat{x}(\omega)$ is close to the empirically observed spectrum of the earthquake-related ground motion, and
- finally, we apply the inverse Fourier transform to $\widehat{m}(\omega)$; the result $m(t)$ of this application is the desired simulation.

It turns out (see, e.g., [2, 5, 6, 11]) that the most accurate simulation results are obtained if we use the window of the type $t^a \cdot \exp(b \cdot t)$, a general type for which the generalized harmonic wavelets are a particular case corresponding to $a = -1$ and imaginary values b.

Problem: this empirical fact needs a theoretical explanation. While there is an empirical evidence of the efficiency of the generalized harmonic wavelets in engineering and seismic applications, to the best of our knowledge, there is no convincing theoretical explanation for this empirical efficiency. Without such a theoretical explanation, it is not clear whether this efficiency is indeed a general phenomenon – or it is simply an artifact of the current applications, and in other applications, other wavelets will turn out to be more efficient.

It is also not clear whether these are indeed the most efficient wavelets. In each empirical study, we can only compare finitely many types of wavelets. As a result, after such comparison, it is not clear whether the selected wavelets are indeed the best of all possible ones – or they are simply the best of all that we tried, and some other not-yet-tried wavelets will be even more efficient.

What we do in this paper. To resolve all these doubts, it is desirable to undertake a theoretical analysis of the corresponding problem. This is what we do in this paper:

- we formulate the problem of selecting the most efficient family of wavelets as a precise mathematical problem, and
- we show that the generalized harmonic wavelets are indeed the most efficient – in this precise sense.

Thus, we indeed provide a theoretical explanation for the efficiency of harmonic wavelets in engineering and seismic applications.

2 Analysis of the Problem and the Resulting Justification

What we need. We are interested in representing how the signal x depends on time t. There may be many different types of signals. In general, we need infinitely many parameters to describe a general signal. However, in a computer, we can only represent finitely many numbers. Thus, a reasonable idea is to consider a finite-parameter family of functions, i.e., to consider functions of the type $x(t) = \sum_{i=1}^{n} c_i \cdot e_i(t)$ for some functions $e_1(t)$, ..., $e_n(t)$. The question is then: which functions $e_i(t)$ should we choose?

Scale- and shift-invariance: idea. The numerical value of time depends on the selection of a measuring unit and on the selection of the starting point.

If we replace the original unit for measuring time with a new measuring unit which is a times smaller, then all the numerical values of time will be multiplied by a: $t \to a \cdot t$. For example, if we replace days with hours, then 2 days becomes $24 \cdot 2 = 48$ hours.

If we replace the original starting point for measuring time with a new starting point which is b units earlier, then the value b is added to all the numerical values of time: $t \to t + b$.

The selection of the measuring unit and the selection of the starting point change the numerical values, but they do not change the physics of the process. It is therefore reasonable to require that the basic functions $e_i(t)$ do not depend on these selections, i.e., that if we change the measuring unit or the starting point, we will still get the same basic function.

Of course, to make sure that all the physical formulas remain valid, we probably need to also accordingly change the related units. For example, distance d is velocity v times time t. This formula remains valid no matter what units we choose for measuring distance and time, but for it to remain valid, we need to appropriately change the unit for measuring velocity.

In our case, for the formula $x(t) \approx e_i(t)$ to remain valid, we may need to change the unit for measuring the signal x. Under this change, the original numerical value of x changes to $C \cdot x$, for some appropriate value C. In other words, we arrive at the following two requirements:

- for every $a > 0$, there exists a value $C(a)$ (depending on a) for which

$$e_i(a \cdot t) = C(a) \cdot e_i(t);$$

  and

- for every b, there exists a value $C(b)$ (depending on b) for which

$$e_i(t + b) = C(b) \cdot e_i(t).$$

The first property is usually called *scale-invariance*, the second *shift-invariance*.

Problem: a function can be scale-invariant or shift-invariant, but not both. Both for scale-invariance and for shift-invariance, the solutions to the above requirements are well-known. Namely, for the case when the signal is measurable (and all physical signals are) [1]:

- scale-invariant signals have the form $x(t) = A \cdot t^\alpha$, for some A and α, and
- shift-invariant signals have the form $x(t) = B \cdot \exp(\beta \cdot t)$, for some B and β.

Comment. In the general case, the proofs of these results are somewhat complicated, but for the case of smooth (differentiable) functions, the proofs are reasonably straightforward. For reader's convenience, we reproduce these proofs in the special proofs section at the end of this paper.

As we can easily see, the only functions that belong to both classes are constant functions. A non-constant function cannot be both scale- and shift-invariant. So what shall we do?

Idea: combining scale- and shift-invariant functions. We can have scale-invariant functions – but this will ignore the requirement of shift-invariance. We can have shift-invariant functions – but this will ignore the requirement of scale-invariance. A more reasonable idea is to consider combinations of scale- and shift-invariant functions, i.e., functions of the type $x(t) = F(c(t), h(t))$, where the function $c(t)$ is s*c*ale-invariant, the function $h(t)$ is s*h*ift-invariant, and $F(c, h)$ is an appropriate combination function.

Which combination function should we choose. As we have mentioned earlier, the numerical value of the signal changes if we change the measuring unit – in which case all the numerical values are multiplied by the same constant. From this viewpoint:

- for every $\lambda > 0$, the function $c(t)$ describes the same signal as the function $\lambda \cdot c(t)$, and
- for every $\mu > 0$, the function $h(t)$ describes the same signal as the function $\mu \cdot h(t)$.

It is reasonable to require that when we make these changes, the combined function remains the same – maybe modulo an appropriate re-scaling of the resulting signal. In other words, we require that for all possible values $c(t)$, $h(t)$, $\lambda > 0$, and $\mu > 0$, we have

$$F(\lambda \cdot c(t), \mu \cdot h(t)) = C(\lambda, \mu) \cdot F(c(t), \mu(t))$$

for some value $C(\lambda, \mu)$ depending on λ and μ.

This functional equation also has a known solution [1]: namely, all its solutions have the form $F(c, h) = D \cdot c^p \cdot h^q$ for some D, p, and q. (For the smooth case, the proof is given in the last section.) Thus, the desired combined functions $e_i(t)$ should have the form

$$e_i(t) = F(c(t), h(t)) = D \cdot (c(t))^p \cdot (h(t))^q,$$

where $c(t)$ is a scale-invariant function and $h(t)$ is a shift-invariant function.

Into this formula, we can substitute the known expression $c(t) = A \cdot t^\alpha$ for scale-invariant functions and the known expression $h(t) = B \cdot \exp(\beta \cdot t)$ for shift-invariant functions. As a result, we arrive at the following formula.

Resulting formula. We conclude that

$$e_i(t) = A_0 \cdot t^{a_0} \cdot \exp(b_0 \cdot t),$$

where $A_0 \stackrel{\text{def}}{=} D \cdot A^p \cdot B^q$, $a_0 \stackrel{\text{def}}{=} p \cdot \alpha$, and $b_0 \stackrel{\text{def}}{=} q \cdot \beta$.

Thus, we have indeed explained the empirically successful formulas.

3 Proofs

In this section, as promised, we show – in the case of smooth (differentiable) signals – that scale- and shift-invariance indeed leads to the desired formulas.

Case of scale-invariance. Let us assume that the differentiable function $x(t)$ satisfies the scale-invariance requirements, i.e.,

$$x(a \cdot t) = C(a) \cdot x(t).$$

The function $C(a)$ is a ratio of two differentiable functions $x(a \cdot t)$ and $x(t)$ and is, thus, differentiable itself.

Since all three terms in the above equality are differentiable, we can differentiate both sides with respect to a. As a result, we get

$$t \cdot x'(a \cdot t) = C'(a) \cdot x(t),$$

where $x'(t)$, as usual, denotes the derivative. In particular, for $a = 1$, we get $t \cdot x'(t) = \alpha \cdot x(t)$, where we denoted $\alpha \stackrel{\text{def}}{=} C'(1)$. This means that

$$t \cdot \frac{dx}{dt} = \alpha \cdot x.$$

We can separate the variables x and t if we multiply both sides by dt and divide both sides by $t \cdot x$. Then, we get

$$\frac{dx}{x} = \alpha \cdot \frac{dt}{t}.$$

Integrating both sides, we get $\ln(x) = \alpha \cdot \ln(t) + C_0$, where C_0 is an integration constant. Thus, by taking exp of both sides, we conclude that $x(t) = A \cdot t^\alpha$, where we denoted $A \stackrel{\text{def}}{=} \exp(C_0)$. The statement is proven.

Case of shift-invariance. Let us assume that the differentiable function $x(t)$ satisfies the shift-invariance requirements, i.e.,

$$x(t + b) = C(b) \cdot x(t).$$

The function $C(b)$ is a ratio of two differentiable functions $x(t+b)$ and $x(t)$ and is, thus, differentiable itself.

Since all three terms in the above equality are differentiable, we can differentiate both sides with respect to b. As a result, we get $x'(t+b) = C'(b) \cdot x(t)$. In particular, for $b = 0$, we get $x'(t) = \beta \cdot x(t)$, where we denoted $\beta \stackrel{\text{def}}{=} C'(0)$. This means that

$$\frac{dx}{dt} = \beta \cdot x.$$

We can separate the variables x and t if we multiply both sides by dt and divide both sides by x. Then, we get

$$\frac{dx}{x} = \beta \cdot dt.$$

Integrating both sides, we get $\ln(x) = \beta \cdot t + C_0$, where C_0 is an integration constant. Thus, by taking exp of both sides, we conclude that $x(t) = B \cdot \exp(\beta \cdot t)$, where we denoted $B \stackrel{\text{def}}{=} \exp(C_0)$. The statement is proven.

Scale-invariant combination functions. We want to find functions for which $F(\lambda \cdot c, \mu \cdot h) = C(\lambda, \mu) \cdot F(c, h)$. In particular, for $\mu = 1$, we have $F(\lambda \cdot c, h) = C_1(\lambda) \cdot F(c, h)$, where we denoted $C_1(\lambda) \stackrel{\text{def}}{=} C(\lambda, 1)$. Specifically, if we fix h, then we get

$$F_h(\lambda \cdot c) = C_1(\lambda) \cdot F_h(c),$$

where we denoted $F_h(c) \stackrel{\text{def}}{=} F(c, h)$.

The above equality for $F_h(c)$ is exactly the scale-invariance requirement that we considered earlier, so we know that $F_h(c) = A \cdot c^p$ for some A and p. Thus, we conclude that

$$C_1(\lambda) = \frac{F_h(\lambda \cdot c)}{F_h(c)} = \frac{A \cdot (\lambda \cdot c)^p}{A \cdot c^p} = \lambda^p.$$

Therefore, the formula $F(\lambda \cdot c, h) = C_1(\lambda) \cdot F(c, h)$ takes the form $F(\lambda \cdot c, h) = \lambda^p \cdot F(c, h)$. In particular, for $c = 1$, we conclude that $F(\lambda, h) = \lambda^p \cdot F(1, h)$, i.e., that $F(c, h) = c^p \cdot F(1, h)$.

Similarly, we can conclude that $F(c, \mu \cdot h) = \mu^q \cdot F(c, h)$, for some q. In particular, for $c = h = 1$, we conclude that $F(1, \mu) = \mu^q \cdot D$, where we denoted $D \stackrel{\text{def}}{=} F(1, 1)$. So, $F(1, h) = D \cdot h^q$. From $F(c, h) = c^p \cdot F(1, h)$, we can now conclude that indeed $F(c, h) = D \cdot c^p \cdot h^q$. The statement is proven.

Acknowledgments

This work was supported in part by the US National Science Foundation grant HRD-1242122 (Cyber-ShARE Center of Excellence).

References

[1] J. Aczél and J. Dhombres, *Functional Equations in Several Variables*, Cambridge University Press, 2008.

[2] G. M. Atkinson and P. G. Somerville, "Calibration of time history simulation models", *Bulletin of the Seismological Society of America*, 1994, Vol. 84, No. 2, pp. 400–414.

[3] M. Beer, I. A. Kouguoumtzoglou, and E. Patelli, "Emerging concepts and approaches for efficient and realistic uncertainty quantification", In: D. Frangopol and Y. Tsompanakis (eds.), *Maintenance and Safety of Aging Infrastructure*, CRC Press/Balkema, Boca Raton, Florida, 2014, pp. 121–161.

[4] M. Beer, I. A. Kouguoumtzoglou, and E. Patelli, "Emerging concepts and approaches for efficient and realistic uncertainty quantification", In: S. Freitag, R. L. Muhanna, and R. L. Mullen (eds.), *Proceedings of the 7th International Workshop on Reliable Engineering Computing*, Bochum, Germany, June 15–17, 2016, pp. 1–39.

[5] D. M. Boore, "Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra", *Bulletin of the Seismological Society of America*, 1983, Vol. 73, No. 6, pp. 1865–1894.

[6] D. M. Boore, "Simulation of groud motion using the stochastic method", *Pure and Applied Geophysics*, 2003, Vol. 160, pp. 635–676.

[7] I. A. Kouguoumtzoglou and P. Spanos, "An analytical Wiener path integral technique for non-stationary response determination of nonlinear oscillators", *Probabilsitic Engineering Mechanics*, 2012, Vol. 28, pp. 125–131.

[8] D. E. Newland, "Harmonic wavelet analysis", *Proceedings of the Royal Society of London, Part A*, 1993, Vol. 443, No. 1917, pp. 203–225.

[9] D. E. Newland, *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*, Longmans Scientific and Technical, Harlow, UK, 1993.

[10] D. E. Newland, "Harmonic and musical wavelets", *Proceeedings of the Royal Society of London, Part A*, 1994, Vol. 444, pp. 605–620.

[11] G. R. Saragoni and G. C. Hart, "Simulation of artifical earthquakes", *Earthquake Engineering and Structural Dynamics*, 1974, Vol. 2, pp. 249–267.

[12] P. Spanos and I. A. Kouguoumtzoglou, "Harmonic wavelet based statistical linearization for response evolutionary power spectrum determination", *Probabilistic Engineering Mechanics*, 2012, Vol. 27, No. 1, pp. 57–68.

[13] P. Spanos, J. Tezcan, and P. Tratskas, "Stochastic processes evolutionary spectrum estimation via harmonic wavelets", *Computer methods in Applied Mechanics and Engineering*, 2005, Vol. 194, No. 12–16, pp. 1367–1383.

[14] R. Yan and R. X. Gao, "Generalized harmonic wavelet as an adaptive filter for machine health diagnosis", In: M. Tomizuka (ed.), *Smart Structures and Materials 2005: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, Proceedings of SPIE, Vol. 5765, pp. 786–793.