Applied Mathematical Sciences, Vol. x, 200x, no. xx, xxx - xxx

MATHEMATICAL JUSTIFICATION OF SPECTRAL/COVARIANCE TECHNIQUES: ON THE EXAMPLE OF ARC DETECTION

Jan Beck and David Nemir

X-L Synergy
P.O. Box 23189
El Paso, TX 79923-3189, USA
email appliedmath@janbeck.com

Vladik Kreinovich

Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
email vladik@utep.edu

Abstract

Detecting arcing faults is an important but difficult-to-solve practical problem. Many existing methods of arc detection are based upon acquiring a signal that is proportional to current and then making an analysis of the signal's power spectrum (or, equivalently, its covariance function). Since the power spectrum, i.e., the *absolute values* of the Fourier transform, carries only *partial* information about the signal, a natural question is: why should we restrict ourselves to the use of this partial information? A related question is caused by the fact that even the most efficient methods still miss some arcing faults and/or lead to false detection; what methods should we use to improve the quality of arc detection?

Our analysis is much more general than the arc detection problem and can be used to justify and select detection methods in other applied problems as well.

Mathematics Subject Classification: 42A38, 42B10

Keywords: arcing faults, Fourier transform, power spectrum

1 Introduction

Electrical systems sometimes start an unplanned "arcing", i.e., producing an electric connection in the normally nonconducting media (e.g., in the air). Arcing can be a result of an unplanned connection between two wires, or the result of an electric wire break down or a connector becoming loose.

Unplanned arcing not only disrupts the normal functioning of an electric system, it can also produce damage. Arcing damage is especially dangerous in aerospace systems where an arc fault can cause the damage to the wiring system causing an aircraft to crash. Because of this danger, it is extremely important to be able to detect the arcing based on the observed electric current or its rate of change; see, e.g., [9]. The corresponding signal will be denoted by $x(t)$.

Arcing is very difficult to detect because every individual arc is different. The material makeup of the wires and of the insulation, air pressure, and arc severity all affect the arc behavior. As a result, arcs are difficult to model, difficult to predict, and difficult to detect.

The problem of detecting unplanned arcs is made even more complex by the fact that in some practical systems such as brush motors, arc welders, and arc discharge lamps, there are permissible arcs. Electric switches and relays are also sources of permissible arcs. For such systems, we must be able to distinguish between the effect of permissible and unplanned arcs.

Many existing methods of arc detection are based on the analysis of the signal's power spectrum $P(\omega) = |\hat{x}(\omega)|^2$, where $\hat{x}(\omega)$ is the Fourier transform of the signal $x(t)$; see, e.g., [4, 7, 8, 10]. Some of the arc detection methods use the covariance function $A(\tau) = \int x(t) \cdot x(t - \tau)\, dt$.

For example, [2] proposed to use the fact that without the arc, a typical electrical system may be characterized by linear equations while with an arc, this is not possible. So, to detects the arc faults, [2] proposes to find the coefficients of the corresponding linear equations by using the Least Squares fit – and use the size of the remainder to detect arcs. The Least Square fit for a signal naturally involves the use of covariance.

While computationally the power spectrum and covariance function are different, from the mathematical viewpoint, they represent the same information in the sense that they can be easily obtained from each other (see, e.g., [3]):

- the power spectrum is equal to the Fourier transform of the covariance function, while

- the covariance function can be obtained from the power spectrum by using the inverse Fourier transform.

Since the power spectrum, i.e., the *absolute values* of the Fourier transform, carries only *partial* information about the signal, a natural question is: why should we restrict ourselves to the use of this partial information? A related question is caused by the fact that even the most efficient methods still miss some arcing faults and/or lead to false detection; what methods should we use to improve the quality of arc detection?

Our analysis is much more general than the arc detection problem and can be used to justify and select detection methods in other applied problems as well.

Our objective is to be understood both by engineers and by mathematicians. As a result, we start with explaining the corresponding engineering problem – this part can skipped by an interested engineer. Then, we explain our mathematical approach in terms that should be understandable to a working engineer; an interested mathematician can easily skip the derivation details.

2 Towards a General Method of Event Detection

We need to be able, given a signal $x(t)$, i.e., a function, to detect a certain event. In our case study, this event is the presence of an arc, or, for a system with permissible arcs, the presence of an unplanned arc.

How can we detect a generic event like that? Since we do not have a good theoretical model for such an event, we have to first gather the signals corresponding to the "normal" situation (where the event did not occur) and signals corresponding to "abnormal" situation – where this event actually occurred. There are also borderline situations in which we are not sure whether an event occurred.

Observed normal signals can be viewed as samples from the set N of all the signals corresponding to the normal situations, while observed abnormal signals can be viewed as samples from the set A of all the signals corresponding to the abnormal situations.

We must be able, given a signal $x(t)$ from the set S of all possible signals, to distinguish whether the signal belongs to the set N or to the set A.

It is reasonable to assume that the sets A and N are, in general, disjoint (= non-intersecting) closed sets in the space S of all possible signals. It is also reasonable to assume that every two such sets can be contained in disjoint open neighborhoods. Thus, we can conclude that there exists a function $J : S \to \mathbb{R}$

and a number $\varepsilon > 0$ for which $J(x) \geq \varepsilon$ for $x \in A$ and $J(x) \leq -\varepsilon$ for $x \in N$; see, e.g., [6].

We want to minimize the number of situations when we cannot tell whether the event occurred or not. Thus, we would like to make the gap between the normal and abnormal signals as narrow as possible. Ideally, we should therefore aim for a function $J : S \to \mathbb{R}$ for which $J(x) > 0$ for $x \in A$ and $J(x) < 0$ for $x \in N$. In these terms, the question is to find an appropriate function $J : S \to \mathbb{R}$.

How can we describe such arbitrary functions?

In principle, there exist events which drastically change the signal; detecting such events is usually reasonably easy. The problem of event detection becomes difficult when the signals corresponding to the abnormal events are very similar to the signals corresponding to the normal events. So, for event detection purposes, it is sufficient to be able to distinguish between close signals, i.e., signals from a small neighborhood. Within this small neighborhood, we can expand the function J in Taylor series and keep only a few terms in this expansion. A signal $x(t)$ is usually represented by its values $x(t_0)$, $x(t_1)$, ..., $x(t_k)$, ..., at different moments of time t_0, $t_1 = t_0 + \Delta t$, $t_2 = t_1 + \Delta t$, ... Thus, a general function $J(x)$ can be represented as a general linear (quadratic, cubic, etc.) function of these variables:

$$J(x) = a_0 + \sum_i a_i \cdot x(t_i) + \sum_i \sum_j a_{ij} \cdot x(t_i) \cdot x(t_j) + \ldots$$

This expression can be further simplified if we take into account that, e.g., the first sum $\sum\limits_i a_i \cdot x(t_i)$ in this expression is an integral sum for the corresponding integral $\int a(t) \cdot x(t)\, dt$. In other words, taking into account the discretization of the signal, it is, in effect, this integral. Similarly, the second sum is an integral sum for the corresponding 2-D integral, etc. As a result, we arrive at the following general expression for $J(x)$:

$$J(x(t)) = a_0 + \int a_1(t) \cdot x(t)\, \mathrm{d}t + \int a_2(t,s) \cdot x(t) \cdot x(s)\, \mathrm{d}t\, \mathrm{d}s + \ldots \quad (1)$$

In general, the expressions $a_1(t)$ and $a_2(t,s)$ do not have to be functions in the normal mathematical sense, we can have "generalized" functions such as delta-functions $\delta(x)$; see, e.g., [1, 5] (In mathematical physics, such "generalized" functions are sometimes called *distributions* – not to be confused with probability distributions.)

It is reasonable to start with the smallest possible number of terms in the Taylor expansion, and then, if needed, add further terms.

An additional feature of most event detection situations is related to the fact that the mathematical representation $x(t)$ of the signal depends on the choice of the starting moment for measuring time. If, instead of the original starting moment $t = 0$, we choose a new moment which is t_0 seconds later, then each moment of time described by the value t on the old scale will now be described by the value $t' = t - t_0$ in the new scale, and the value t' in the new scale corresponds to the value $t = t' + t_0$ on the old time scale. Thus, the value $x'(t')$ of the signal in the new time scale corresponds to the value $x(t' + t_0)$ in the old scale: $x'(t') = x(t' + t_0)$. In other words, in the new time scale, instead of the original function $x(t)$, we have a new function $x(t' + t_0)$ (representing the same signal). In most practical problems (including the problem of arc detection), there is nothing special about the moment $t = 0$, so we should expect that the corresponding functional does not change if we simply change the time origin: $J(x(t)) = J(x(t' + t_0))$.

We will see that this time-invariance leads, in effect, to the justification of the spectral techniques.

3 First Auxiliary Result: Linear Functionals are Not Sufficient

As we have mentioned earlier, it is reasonable to start with the Taylor expansion with the smallest possible number of terms, i.e., with a linear expression

$$J(x(t)) = a_0 + \int a_1(t) \cdot x(t)\,\mathrm{d}t. \tag{2}$$

For this expression, time-invariance $J(x(t)) = J(x(t+t_0))$ means that for every t_0 and for every signal $x(t)$, we have

$$a_0 + \int a_1(t) \cdot x(t)\,\mathrm{d}t = a_0 + \int a_1(t) \cdot x(t + t_0)\,\mathrm{d}t. \tag{3}$$

By introducing a new variable $t' = t + t_0$ (for which $t = t' - t_0$ and $\mathrm{d}t = \mathrm{d}t'$), we can transform the integral in the right-hand side into

$$\int a_1(t' - t_0) \cdot x(t')\,\mathrm{d}t'.$$

Renaming t' into t and canceling the common term a_0 in both sides, we conclude that with a linear expression for J,

$$\int a_1(t) \cdot x(t)\,\mathrm{d}t = \int a_1(t - t_0) \cdot x(t)\,\mathrm{d}t, \tag{4}$$

i.e., that

$$\int (a_1(t) - a_1(t - t_0)) \cdot x(t)\,\mathrm{d}t = 0 \tag{5}$$

for all possible signals $x(t)$. Since this linear function (integral) is 0 for all the inputs $x(t)$, this means that all the coefficients at this linear expression must be equal to 0, i.e., that we should have $a_1(t) = a_1(t - t_0)$ for all t and all t_0. In particular, for $t_0 = t$, we conclude that $a_1(t) = a_1(0)$, i.e., that $a_1(t)$ is a constant: $a_1(t) = a_1 \stackrel{\text{def}}{=} a_1(0)$ for all t. Thus, the general time-invariant linear expression $J(x)$ takes the form

$$J(x(t)) = a_0 + a_1 \cdot \int x(t)\,\mathrm{d}t.$$

This expression depends only on the average value $\int x(t)\,\mathrm{d}t$ of the signal $x(t)$. The time average of the signal $x(t)$ is often zero. For example, in most arc sensing systems, $x(t)$ is often obtained from a current sense transformer (see, e.g., [2]), that removes the direct current component. As a result, based on the time average of the signal, we usually cannot tell whether an event occurred or not. Definitely we cannot tell based on the time average whether there is arcing or not.

Thus, linear terms are not enough, so quadratic (or maybe even higher) terms are needed.

4 Main Result: Quadratic Time-Invariant Functionals are (Almost) Uniquely Determined by the Power Spectrum

For the general quadratic functional

$$J(x(t)) = a_0 + \int a_1(t) \cdot x(t)\,\mathrm{d}t + \int a_2(t,s) \cdot x(t) \cdot x(s)\,\mathrm{d}t\,\mathrm{d}s, \qquad (6)$$

time-invariance $J(x(t)) = J(x(t + t_0))$ means that for every t_0 and for every signal $x(t)$, we have

$$a_0 + \int a_1(t) \cdot x(t)\,\mathrm{d}t + \int a_2(t,s) \cdot x(t) \cdot x(s)\,\mathrm{d}t\,\mathrm{d}s =$$

$$a_0 + \int a_1(t) \cdot x(t+t_0)\,\mathrm{d}t + \int a_2(t,s) \cdot x(t+t_0) \cdot x(s+t_0)\,\mathrm{d}t\,\mathrm{d}s. \qquad (7)$$

By introducing new variables $t' = t + t_0$ and $s' = s + t_0$ (for which $t = t' - t_0$, $s = s' - t_0$, $\mathrm{d}t = \mathrm{d}t'$ and $\mathrm{d}s = \mathrm{d}s'$) and canceling the common term a_0 in both sides, we conclude that

$$\int a_1(t) \cdot x(t)\,\mathrm{d}t + \int a_2(t,s) \cdot x(t) \cdot x(s)\,\mathrm{d}t\,\mathrm{d}s =$$

$$\int a_1(t-t_0)\cdot x(t+t_0)\,\mathrm{d}t + \int a_2(t-t_0,s-t_0)\cdot x(t+t_0)\cdot x(s+t_0)\,\mathrm{d}t\,\mathrm{d}s, \quad (8)$$

i.e., that

$$\int (a_1(t)-a_1(t-t_0))\cdot x(t)\,\mathrm{d}t +$$
$$\int (a_2(t,s)-a_2(t-t_0,s-t_0))\cdot x(t)\cdot x(s)\,\mathrm{d}t\,\mathrm{d}s = 0 \quad (9)$$

for all possible signals $x(t)$. Since this quadratic function (integral) is 0 for all the inputs $x(t)$, this means that all the coefficients at this quadratic expression must be equal to 0, i.e., that we should have $a_1(t)=a_1(t-t_0)$ for all t and all t_0, and that we should have $a_2(t,s)=a_2(t-t_0,s-t_0)$ for all t, s, and t_0. We already know that the condition $a_1(t)=a_1(t-t_0)$ implies that $a_1(t)$ is a constant: $a_1(t)=a_1$ for all t. Similarly, for the second condition, for $t_0=t$, we get $a_2(t,s)=b(s-t)$, where we denoted $b(t)\stackrel{\text{def}}{=}a_2(0,t)$. Thus, the general time-invariant quadratic functional has the form

$$J(x(t)) = a_0 + a_1\cdot\int x(t)\,\mathrm{d}t + \int b(s-t)\cdot x(t)\cdot x(s)\,\mathrm{d}t\,\mathrm{d}s. \quad (10)$$

By definition of the Fourier transform, the first integral $I_1\stackrel{\text{def}}{=}\int x(t)\,\mathrm{d}t$ is equal to $\hat{x}(0)$. This value is always a real number and thus, it can be almost uniquely determined by the corresponding value $|\hat{x}(0)|^2$ of the power spectrum – with the only uncertainty is that from the power spectrum, we cannot determine the sign of this integral.

The second integral $I_2\stackrel{\text{def}}{=}\int b(s-t)\cdot x(t)\cdot x(s)\,\mathrm{d}t\,\mathrm{d}s$ can be described as $I_2=\int y(s)\cdot x(s)\,\mathrm{d}s$, where $y(s)\stackrel{\text{def}}{=}\int b(s-t)\cdot x(t)\,\mathrm{d}t$ is a *convolution* of the functions $b(t)$ and $x(t)$. The Fourier transform $\hat{y}(\omega)$ of y is thus equal to the product of the Fourier transforms: $\hat{y}(\omega)=\hat{b}(\omega)\cdot\hat{x}(\omega)$.

Due to the Parseval's (Plancherel) Theorem (see, e.g., [3]),

$$I_2 = \int y(s)\cdot x(s)\,\mathrm{d}s = \int \hat{y}(\omega)\cdot\hat{x}^*(\omega)\,\mathrm{d}\omega$$

(where x^*, as usual, means a complex conjugate). Substituting the above expression for $\hat{y}(\omega)$, we conclude that

$$I_2 = \int \hat{b}(\omega)\cdot\hat{x}(\omega)\cdot\hat{x}^*(\omega)\,\mathrm{d}\omega.$$

Since $a\cdot a^* = |a|^2$, we thus get $I_2=\int \hat{b}(\omega)\cdot|\hat{x}(\omega)|^2\,\mathrm{d}\omega$, so we conclude that

$$J(x(t)) = a_0 + a_1\cdot\hat{x}(0) + a_2\cdot\int \hat{b}(\omega)\cdot|\hat{x}(\omega)|^2\,\mathrm{d}\omega. \quad (11)$$

In other words, every time-invariant quadratic functional depends only on the spectrum $|\hat{x}(\omega)|^2$ – and on the value $\hat{x}(0)$ which is equal to $\pm\sqrt{|\hat{x}(0)|^2}$.

5 Second Auxiliary Result: Cubic Time-Invariant Functionals are Not Uniquely Determined by the Power Spectrum

In the previous section, we have shown that, in effect, every time-invariant quadratic functional depends only on the power spectrum. One may think that the same is true for higher-order functionals as well, but this is not true already for cubic functionals. For example, let us pick any two positive numbers $\omega_1 \neq \omega_2$ and define the following simple cubic functional:

$$J(x) = \mathrm{Re}[\hat{x}(\omega_1) \cdot \hat{x}(\omega_2) \cdot \hat{x}^*(\omega_1 + \omega_2)], \tag{12}$$

where $\mathrm{Re}(a)$ denotes the real part of the complex number a. The numerical value of this functional is not uniquely determined by the power spectrum $|\hat{x}(\omega)|^2$: indeed, if we retain the absolute value (magnitude) $|\hat{x}(\omega_1)|$ of the complex number $\hat{x}(\omega_1)$ but change the phase, that will change the value of the real part and thus, of the functional.

On the other hand, this functional is time-invariant. Indeed, if we replace $x(t)$ with $x_{\mathrm{new}}(t) \stackrel{\mathrm{def}}{=} x(t + t_0)$, then the new Fourier transform takes the form $\hat{x}_{\mathrm{new}}(\omega) = \hat{x}(\omega) \cdot e^{\mathrm{i} \cdot \omega \cdot t_0}$. Thus, for the new function $x_{\mathrm{new}}(t)$, the functional has the value

$$J(x_{\mathrm{new}}) = \mathrm{Re}[\hat{x}_{\mathrm{new}}(\omega_1) \cdot \hat{x}_{\mathrm{new}}(\omega_2) \cdot \hat{x}^*_{\mathrm{new}}(\omega_1 + \omega_2)], \tag{13}$$

i.e.,

$$J(x_{\mathrm{new}}) = \mathrm{Re}[\hat{x}(\omega_1) \cdot e^{\mathrm{i} \cdot \omega_1 \cdot t_0} \cdot \hat{x}(\omega_2) \cdot e^{\mathrm{i} \cdot \omega_2 \cdot t_0} \cdot \hat{x}^*(\omega_1 + \omega_2) \cdot e^{-\mathrm{i} \cdot (\omega_1 + \omega_2) \cdot t_0}]. \tag{14}$$

Since

$$e^{\mathrm{i} \cdot \omega_1 \cdot t_0} \cdot e^{\mathrm{i} \cdot \omega_2 \cdot t_0} \cdot e^{-\mathrm{i} \cdot (\omega_1 + \omega_2) \cdot t_0} = 1,$$

we thus conclude that $J(x_{\mathrm{new}}) = J(x)$, i.e., that the above cubic functional $J(x)$ is indeed time-invariant.

6 What This Means for a Practitioner

From an engineering viewpoint, we observe that most existing techniques for arc detection (and similar detection problems) use quadratic detection criteria. It turns out that all these methods, in effect, belong to the same class of methods – method which use the signal's power spectrum for detection. Our conclusion is that if we want to further increase the efficiency of the detection techniques, we need to use higher-order methods.

ACKNOWLEDGEMENTS. This work was supported in part by NSF grant HRD-0734825 and by Grant 1 T36 GM078000-01 from the National Institutes of Health.

References

[1] G. Arfken, *Mathematical Methods for Physicists,* Orlando, Florida, Academic Press, 1985.

[2] J. Beck and D.C. Nemir, Arc Fault Detection through Model Reference Estimation, Proc. 2006 Aerospace Congress, *Proceedings of the Society Of Automotive Engineers (SAE) Power Systems Conference*, New Orleans, Louisiana, SAE paper #2006-01-3090.

[3] R. Bracewell, *The Fourier Transform and Its Applications*, McGraw-Hill, New York, 1999.

[4] S.J. Brooks, J.W. Dickens, and W.H. Strader, *Arcing Fault Detection System*, U.S. Patent 5,682,101, October 28, 1997.

[5] I.M. Gelfand and G.E. Shilov, *Generalized Functions,* Academic Press, New York and London, 1977.

[6] J.L. Kelley, *General Topology,* Springer, New York, 1975.

[7] P. Meckler, K.J. Eichhorn, and W. Ho, Detecting and extinguishing of arcs in aircraft electrical systems, *Proc. of the 2001 Aerospace Congress*, Seattle, Washington, September 10–14, 2001, paper # SAE-2001-01-2657.

[8] B.D. Russell and B.M. Aucoin, *Arc Spectral Analysis System*, U.S. Patent 5,578,931, November 26, 1996.

[9] Underwriters Laboratories, *UL1699 Standard for Arc Fault Circuit Interrupters*, April 2006.

[10] J.J. Zuercher and C.J. Tennies, *Arc Detection Using Current Variation*, U.S. Patent 5,452,223, September 19, 1995.

Received: Month xx, 200x